

Unreasonable hope for Rio fades

The prospect that the 'Earth summit' planned for June at Rio de Janeiro will be a monumental disappointment seems to have been strengthened by the planning meeting in New York over the past five weeks.

Just how the United Nations (UN) Conference on Environment and Development came to be known as the 'Earth Summit' is not a pointless question, for its answer should identify the means by which have been raised expectations so great that they are bound to be frustrated. When the conference was first sanctioned by the UN Assembly, there were several practical tasks to be tackled. Could there be some kind of understanding about relations between industrialized and developing countries on environmental matters? Should development assistance through organizations such as the World Bank be tied to stipulations about environmental good behaviour, for example, and how should developing countries be compensated for the extra costs entailed? Then came renewed anxiety about the threat of global warming and the assumption that the meeting at Rio would be the occasion for the solution (at least on paper) of all the world's environmental problems. Now it is only a matter of time, and a quite short time at that, before the more joyous expectations are disappointed.

So much is clear from reactions to the meeting in New York that has occupied the past five weeks, at which government representatives have been attempting to hammer out an agenda for Rio. The issue has not been so much that of what to talk about, but of the terms in which it should be discussed. Those looking for far-reaching agreements at Rio have been disappointed that little has yet been agreed about the limits on the emission of carbon dioxide and other greenhouse gases, let alone about the terms on which developing countries would be compensated for maintaining the diversity of their natural species and the extent of their rain forests. But the expectation was always unrealistic, and there is a sense in which the expressions of disappointment are feigned, and are therefore false.

That does not imply that there is no global environmental problem. So much should be clear from statements such as that put out some weeks ago by the Royal Society of London and the US National Academy of Sciences (see *Nature* 355, 752; 1992), not to mention the army of earlier commentators on the subject. But it is not one problem, but several. The difficulty, which the organizers of the Earth summit appear deliberately to have obscured, is that of assigning some degree of priority to them. Even at this stage, is it necessarily too late?

So far, two strictly global problems of an environmental character have been identified, of which the most urgent is the threat of global warming caused primarily by the

accumulation of carbon dioxide. (Ozone depletion caused by chlorofluorocarbons [CFCs] has already been demonstrated, but its direct consequences, so far as is known, are less serious.) Given the inherent uncertainties, there is no way of telling exactly what the degree of climatic change will be, or when the pace of climatic change will exceed the capacity of at least industrialized communities on the surface of the Earth to adapt to change, but the chance that global warming is a mirage diminishes as the seasons pass. And the end-point, perhaps a century or perhaps several centuries away, is the melting of the surviving ice caps in the Antarctic and in Greenland, with the calamitous consequences that would bring.

Those are the reasons, often repeated in opinions such as these, why global warming should be at the centre of attention at Rio. To cloud that issue with other questions is a mistake, but one that is probably now irreparable. And there is no question that the containment of global warming (or even a decision by the chief emitters of greenhouse gases not to bother) entails questions of equity between industrialized and developing countries. Undertakings to contain greenhouse gas emissions to the level reached in some previous year would be quickly negated were the emission of carbon dioxide by industrializing states quickly to increase. So an attempt to head off global warming must be a global undertaking, requiring a compact between the chief polluters and the developing countries. So much is generally agreed.

So why should not everybody troop to Rio to sign a binding convention on the subject? Because not nearly enough is known about the nature of the physical problem, let alone about the basis of a lasting deal between the participants. What, for example, is the equitable way to trade off continued emission of CFCs against carbon dioxide? What allowance should be made, in the calculation of quotas for allowed emissions, for continued population growth in developing countries? And so on. None of these questions could be decided quickly. That is why the most that can be hoped for from Rio, but also the best outcome at this stage, is a joint declaration that global warming is a problem that may well have to be reckoned with, a necessarily incomplete list of the questions that must be decided before a convention can be signed and a definition of a mechanism by which those decisions can be made. To ask for more is to ask for a deal on global warming that would be unsatisfactory in one way or



another, and which for that reason would not last.

Such a modest agenda might even persuade President George Bush to take time out from his election campaign to attend, as the president of Brazil, Fernando Collor de Mello, has been urging. Bush's obvious disinclination is no doubt partly inspired by the fear of being trapped into rejecting demands for specific restraints on fossil fuel consumption in the months before a election. But electoral considerations (and the chance that there might be a few green votes to be picked up) might similarly have him on the way to Rio to lend support to the framework of an undertaking to act responsibly, in the interests of the world's environment, if and when the basis of responsible action has been defined. And the plain truth is that no agreement on global warming will be worth having unless it includes the United States (not to mention China and the states of the Commonwealth of Independent States). It will be a great shame if the opportunity of Rio is lost because the enthusiasts have attempted too much. □

Another election soon

Flurries of interest in science towards the end of Britain's election campaign promise little, but another election may.

THE best thing to say about the British general election campaign that ends today (9 April) with an election is that it may have to be repeated very soon. The polls, which suggest that there will be a change of government, also agree that the new government's majority will be only small, and may not even be absolute. Only next week will it become possible to tell whether the new government has been able to muster enough support to last for a reasonable length of time. Under the British system, there could be another election in just a month, but the chances are that everybody is too bored (and party funds too depleted) for that to be possible. But another election in a few months would at least bring into the open the questions buried during the past few weeks of cautious campaigning.

One of these was science. Only towards the end of a contrived campaign did the issue come to life in any form, and then only because the government was provoked into responding to criticisms by the pressure group "Save British Science" and by various round robins carrying the signatures of researchers and academics. Poor Mr Allan Haworth, the minister at the Department of Education and Science with responsibility for higher education and research, was forced into several repetitions of the government's line that there is nothing wrong with the British research enterprise; how could there be, when the public funds available have been increased faster than inflation, and when Britain has "the best university system in Western Europe"? It is to be hoped that by the time the next election comes around, the complacency of Mr Haworth's party will have been punctured by the now-public testaments of British academics to the demoralization that they share.

Before the next election, it is also important that Britain

should learn more about the plans of the two major parties on technical and vocational education, their common recipe for the improvement of the competitiveness of British industry. (The recipes differ in that the Conservative Party offers training vouchers to students not staying on for higher education, would offer vocational examinations to those still at school and would otherwise work through existing Technical Education Councils; the Labour party would develop a "coherent" programme.) What neither party seems fully to appreciate is that industrial competitiveness also requires a cadre of people educated in the methods of research and capable of fostering radical innovation. The demoralization of the research enterprise brought about in the past dozen years is unlikely to increase the supply of such people. □

Bring back Tsongas

US presidential candidates should either forget science or learn a little, if a calamitous forum last week is any guide.

IF only US science could be simply ignored by political candidates, as British science has been. If there had been no mention of it by the presidential hopefuls, it might have been possible to imagine that each cares a little about research, but realizes that there are no votes to be won from common folk by touting it on the campaign trail. But, as luck will have it, the silence — ignorant or otherwise — has not lasted.

Last week, the Council on Research and Technology (Coretech) asked the presidential candidates to describe their positions on science and technology (via representatives) at a special forum in Washington. Three — Mr Jerry Brown, Governor Bill Clinton and President George Bush — complied. For anyone who still thinks that research is at least as important a political issue as bounced checks, the occasion was neither edifying nor reassuring.

Brown was represented by a Baptist reverend from Detroit named Horace Hilsman, who wanted to know why "the AIDS cure in Africa" is being suppressed. Bush's representative, Ms Deborah Wince-Smith, who once campaigned for President Ronald Reagan and who is now assistant secretary for technology policy in the Department of Commerce, let her allegiance show by crediting "President Reagan" with having launched her own agency's National Technology Initiative last year. And Clinton's man momentarily overlooked allegations of scandal against his master when he assured questioners (asking for a position on research issues such as indirect costs and misconduct) that Clinton "is opposed to abuses of all kinds".

This pathetic showing will make for difficult choices at election time in November: vote for Clinton, whose representative embarrassed himself least, or Pat Buchanan, who sent no one at all? If Texas industrialist H. Ross Perot (who has a long track record of research advocacy) enters the race, choice may be simplified. But the best hope may be for Paul Tsongas to re-enter the race. □

US genome head faces charges of conflict

- Watson expected to leave this summer
- Feud with Healy, Bourke led to showdown

Washington

NOBEL Laureate James Watson, the director of the US National Institutes of Health (NIH) genome project, is under investigation for alleged conflict of interest and is expected to resign his position rather than fight what his associates say is a concerted campaign to unseat him.

NIH officials say that Bernadine Healy, the NIH director, has raised concerns about stock that Watson owns in biotechnology and gene sequencing companies that she feels represents a possible conflict of interest. Healy reviewed Watson's files and told the agency's ethics officer, Jack Kress, about a number of such holdings. Kress, who met with Watson on 24 March, believes that Healy could resolve the conflicts by simply signing a waiver to allow Watson to keep the stocks. But Healy "has concerns about signing a waiver" and is reluctant to make such an exemption given the volatile nature of the issue, according to her spokeswoman, Joanna Schneider.

Although Watson could sell the stock in question and resolve the issue, he has told associates that he instead will resign, perhaps as early as 1 August. Contacted at home early this week, Watson said he had been forbidden to talk about the situation and would not comment further.

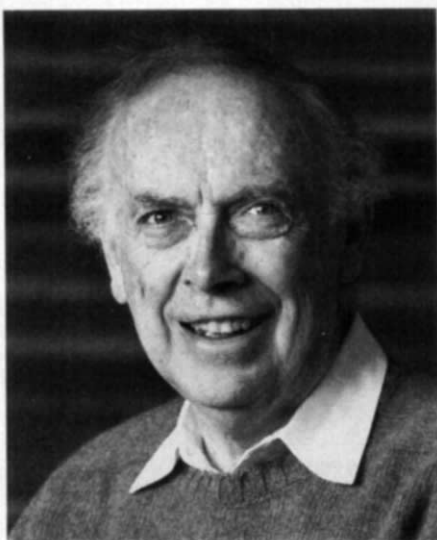
Watson's associates are concerned that the alleged conflict of interest is merely a smokescreen. They believe that last month's showdown marks the latest turn in a running feud between Watson and Healy that has been fanned by Watson's vocal opposition to Healy's decision to file a patent for cDNA sequences last summer (*Nature* **353**, 485; 1991) and his attempts to derail a gene sequencing company proposed by Frederick Bourke, a wealthy entrepreneur (*Nature* **355**, 483; 1992).

Known for his outspoken and often acerbic leadership of the US genome project, Watson's three-year tenure at NIH has been controversial from the start. But in opposing Bourke, who is friendly with several prominent politicians in Washington, including Senate Majority leader George Mitchell, Watson appears to have made one enemy too many.

Bourke has been trying to sign up two prominent gene sequencing researchers — Robert Waterston of the Washington University in St. Louis and John Sulston of the UK Medical Research Council (MRC) Laboratory of Molecular Genetics at Cambridge — as the core of his new company. Watson feared that such a move

might halt one of the most productive collaborations in the genome project and has tried hard to block it.

For example, in February Watson met with MRC officials in Britain to urge them to give Sulston more money to keep him in place. MRC apparently agreed. Sulston has not joined Bourke's company (which is not yet incorporated or named) and the



Watson may resign over charges.

MRC is planning to increase Sulston's funding significantly, although it has not released any details.

While in Britain, Watson also met with officials at Glaxo, the British pharmaceutical company. Accounts of the meeting differ, but Bourke, in a letter last month to Healy and White House officials, alleged that Watson encouraged Glaxo to start its own gene sequencing company centred around Sulston. Such a step would have pre-empted Bourke's plans. In his letter, the details of which have been confirmed by several sources, Bourke also alleged that Watson owns some Glaxo stock and stands to profit should the new venture become a success.

Late last month, after receiving the Bourke letter, Healy passed the allegations on to Kress. "She has a concern about the appearance of, and actual, conflict of interest" in the charges, says NIH's Schneider. "The public deserves to know that there is no conflict of interest — that's Dr. Healy's responsibility."

But Kress disagrees, and says that he found "no substance" to Bourke's charges. The allegations "gave me pause", he says,

"but after talking it over with Dr. Watson, I was satisfied that there was no conflict."

Because Watson continues to serve as director of the Cold Spring Harbor Laboratory, which receives NIH funding, he has already removed himself from the NIH grants process. Although Kress dismissed the allegations contained in the Bourke letter, he determined that some of Watson's other stock holdings might pose a conflict because of the impact of Watson's decisions on the entire biotechnology industry. He says he intended to recommend a waiver after a second meeting with Watson this week, despite Healy's resistance to such a resolution.

Kress confirmed that Watson told him he was planning to resign over the issue. But he emphasized that "there is no ethical reason for him to leave. I am in no way, shape or form recommending that he step down based on anything we discussed."

Watson is said to be treating Healy's ultimatum as a virtual sacking. Schneider says Watson has not been fired. "It's not as sinister as it seems," she says, "but there were some concerns expressed that need to be looked into. Some very difficult decisions are going to have to be made."

Although his associates say that Watson periodically has threatened to resign as a negotiating tactic, they say that this time is different. Norton Zinder, a Rockefeller University geneticist and former chairman of Watson's genome advisory panel, describes the clash as "a political in-fight that has unfortunately reached an impasse."

Earlier this year, NIH officials reviewed Watson's files and found that conflict-of-interest concerns had been raised several times over the past few years. But there is no record in the files of any waiver or other resolution of the issue, according to one official. Bourke's letter appears to have reinforced Healy's resolve to confront Watson on the issue. Watson's "relations with Healy have never been good," says Rich Roberts, Watson's deputy at Cold Spring Harbor, "but the letter from Bourke may have been the last straw."

Researchers are concerned that their feud may damage NIH and the genome project. Coming only days after congressional hearings on the agency's budget, and during a year in which funding is already squeezed by economic concerns, a battle between Watson and Healy could mean a leadership vacuum that costs both the agency and the genome project millions of dollars. Both Zinder and Robert Cook-Deegan, an analyst at the Institute of Medicine and a confidant of Watson, say they have urged Watson to remain on the job until the agency's funding is set, probably sometime in early autumn. But Kress says that Watson "is feeling the crush" and will probably resign sooner than he had planned.

Christopher Anderson

Trying to shake Japan's faith in forecasts

Tokyo

THE future of Japan's huge national research effort to predict earthquakes, which has run for nearly three decades and consumed thousands of millions of yen, is being questioned for the first time by scientists here.

Japan is practically the only country in which earthquake mitigation and hazard reduction takes a back seat to predicting the next big one. Yesterday (8 April), scientists who are critical of that emphasis were expected to have their first chance to modify its course at a meeting in Kyoto of a newly formed subcommittee of the Science Council of Japan, an academic advisory body to the government. But even they are not optimistic about their chances of success.

Standing firm against any attempts to open up the system to outsiders are a handful of senior seismologists, who have been in the earthquake prediction business since the 1960s. This group have effectively by-passed external review by holding a series of closed meetings over the past few months to draw up their own internal review of the research for the government, in much the same way as they have done for the past 30 years. The review process now unfolding provides insight into how small groups of powerful academics can exert enormous influence over government funding of research in Japan.

The earthquake research programme is one of Japan's largest and oldest national research project. In the past 17 years, according to government statistics, the programme has consumed almost 100,000

million yen (US\$700 million), and currently receives about \$50 million a year (see figure). But, the government figures are deceptively low, because they do not include the salaries of the 500 or so researchers in various government research organizations who are involved in earthquake prediction.

Last year, US geophysicist Robert Geller of Tokyo University sent shock waves through the earthquake prediction

In February's issue of the semi-government publication 'Japanese Scientific Monthly' (*Gakujutsu Geppo*), Uyeda describes the growing disillusionment among Japan's seismologists with the present state of prediction research in Japan, which consists mainly of collecting huge amounts of data in the hope that some reliable 'precursor' of earthquakes will be found. That approach has largely been abandoned in the United States and elsewhere after

the failure in the 1970s to find such precursors. Uyeda suggests that perhaps the only way to reinvigorate the research community in Japan is to create "a crisis" and to cut the government budget for the programme to "one-tenth its present size".

The new science council subcommittee, which met for the second time this week in Kyoto, was set up last October partly in response to such criticisms. The 27-man subcommittee includes Geller and several other critics, as well as senior leaders of the earthquake prediction programme. At the urging of the critics, the agenda for this week's meeting calls for a new "blueprint" for earthquake prediction to replace the one that set the present programme on its course in the 1960s.

The discussions in Kyoto, however, are not expected to change the course of earthquake research in Japan. A smaller but much more powerful eight-man committee of senior seismologists, many of whom are also on the new Science Council subcommittee, have met six times behind closed doors since October and have already produced a draft review of the earthquake prediction programme, now in its sixth

Steady funding, but to what end?

All figures in millions

	Earthquake prediction	Volcano prediction
1986	¥ 5,339 (US\$40)	¥ 488 (US\$4)
1987	5,320	595
1988	5,701	620
1989	6,020	674
1990	6,192	731
1991	6,669	698
Total	¥ 35,241 (US\$261)	¥ 3,806 (US\$28)

community in Japan by arguing (*Nature*, **352**, 275; 1991) that earthquake prediction is beyond the present capabilities of science and that Japan's earthquake prediction programme should undergo thorough external review. Following Geller's comments, many Japanese scientists have started to criticise the present system openly. The most recent is Seiya Uyeda, a professor of both Tokai University and Texas A & M University in the United States.

Vulcanologists seek to understand processes

Tokyo

ALTHOUGH Japan's earthquake prediction programme may not have much chance of undergoing significant change (see above), the same is not true for efforts to modify the study of volcanic eruptions.

One difference is that the chairman of the review committee of the volcanic prediction programme, Yoshiaki Ida of the Earthquake Research Institute of Tokyo University, is comparatively young and has strong views about where the field should be headed. He would like to see less emphasis on the empirical collection of data, and more effort spent on trying to understand the basic processes of volcanic eruption. That might include a search for magma chambers and the migration of magma using seismic techniques, as well as modelling such processes. Such an understanding, he believes, could mean a much more systematic approach to prediction rather than the

present way of "just trying to get precursors without physical reason".

Ida would also like to see Japan's vulcanologists spend more time communicating with the public. Japan's volcanic prediction committee has come under severe public criticism in recent years for its terse, uninformative and sometimes misleading statements about volcanic eruptions (*Nature* **351**, 511; 1991). In contrast, Ida says, the Philippine government preceded last year's eruption of the Pinatubo volcano by distributing videos showing the dangers of pyroclastic flows. It also issued public warnings, of increasing severity, in the weeks leading up to the actual eruption.

Ida says he does not expect "drastic" change in Japan's volcanic eruption prediction programme because his review committee is "rather conservative". But he does see hope for a "better trend".

D.S.

Italians first to use stem cells

London

AN Italian research group last week performed the first human gene transfer involving stem cells in the hope of curing a 5-year-old child with a rare genetic disease.

The treatment extends the work of Michael Blaese and French Anderson of the US National Institutes of Health (NIH), who since 1990 have used circulating lymphocytes to treat



Bordignon leads the way

two patients with adenosine deaminase (ADA) deficiency. The US scientists removed the patient's lymphocytes, transfected them with the ADA gene and reinjected them into the patients. But lymphocytes live for only a few

months, whereas stem cells in theory have an unlimited life span.

ADA deficiency is a fatal disease that, at present, can be treated with conjugated bovine ADA or, occasionally, with bone-marrow transplantation. It is extremely rare, a fact that has hindered the use of such new genetic therapy techniques.

The Italian team, led by Claudio Bordignon of the H San Raffaele Research Institute in Milan, received permission in December 1990 from both its institutional ethics committee and the Italian Committee for Biosafety. Its approach is based on the original NIH protocol submitted by Blaese and Anderson, but modified slightly to include manipulation of the cells by a modified retrovirus.

However, a similar protocol from the US group has drawn criticism. After the recent approval by the NIH's Recombinant DNA Advisory Committee (RAC), members of the RAC's human gene therapy subcommittee, which was not consulted, questioned whether the accepted procedure for approval was followed, and worried that retroviruses not fully tested would be introduced into humans. However, the Italian committees were satisfied that Bordignon's amendment did not significantly change the original proposal for lymphocyte manipulation which had been approved in 1990 after extensive review by the NIH and the Food and Drug Administration (FDA).

Bordignon says that his work with paediatrician Alberto Ugazio is an important step forward for gene therapy, both ethically and clinically. First, there is a chance that the graft will take and the child will be completely cured. Second, the technique

makes possible consideration of important questions concerning the long-term survival of both lymphocytes and stem cells. By using two very similar vectors (differing only in one base pair) to transfer the ADA gene into peripheral lymphocytes or into stem cells, the scientists will be able to monitor circulating ADA-positive lymphocytes and determine whether they derived from stem cells.

The new protocol from Blaese and Anderson, which was reviewed in January, was considered to be an amendment to their original proposal, rather than a new proposal (*Nature* 346, 402; 1990). For this reason, the RAC chose not to consult its human gene therapy subcommittee.

Blaese and Anderson want to obtain stem cells from the very low levels circulating in the blood to avoid bone marrow sampling. They plan to prestimulate patients with granulocyte-colony stimulating factor (G-CSF) to increase their abundance. They now await clearance from the FDA and, ultimately, the NIH director.

Like the Italians, they will be using a modified version of their original virus to transfect stem cells so that the origin of surviving ADA-positive cells can be monitored. Dusty Miller of the Fred Hutchinson Cancer Research Center in Seattle, a member of the gene therapy subcommittee with expertise in vectors, believes that the RAC was wrong not to treat the protocol, with the modified as well as the original virus, as new.

"It is no trivial matter to say that using a new retrovirus is okay, just because retroviruses have been used before" says Miller. In other cases, the subcommittee has blocked approval for protocols where the investigator had unknowingly included an oncogene in the viral packaging.

Miller agrees that the Italian protocol poses no extra risk to the patient. But that is no justification for trying to bypass established routes for protocol approval and introducing undue elements of risk into early clinical trials, he says.

Both the American and Italian strategies for transfecting stem cells are simple adaptations of those used for lymphocytes. By contrast, a Dutch research group headed by Dinko Valerio received approval on 20 February for bone-marrow gene therapy based on an entirely different protocol (*Nature* 355, 190; 1992). It would allow them to condition patients fully to receive transformed stem cells from their bone marrow. Valerio says their data from long-term studies on rhesus monkeys have demonstrated the persistence of ADA-positive cells from all haematopoietic lineages. But the group has not yet gone ahead with its trial because there are no ADA patients in Holland.

Alison Abbott

five-year phase. It is this draft, not the deliberations of the science council panel, that will form the basis of government decisions about the next five-year plan. That plan would go into effect in two years.

The drafting committee was appointed by the earthquake prediction subcommittee of the Geodetic Council of Japan, an advisory body to the Ministry of Education, Science and Culture (MESC). Some senior scientists, such as Harumi Aoki of Nagoya University, are members of both groups. And these same senior people will examine the draft and make recommendations to the government this summer in time for the budget requests for fiscal year 1993, which begins next April.

This leaves the new Science Council subcommittee "powerless", says Geller. Geller and other critics have not even been shown the drafting committee's report, and there are no plans for them to see it. Tomowo Hirasawa of Tohoku University, who is chairman of both the new subcommittee and the drafting committee, says the views of the new subcommittee will be conveyed to the MESC earthquake prediction subcommittee by "individual members", in other words, by the same people who drafted the review.

The contents of the draft review remain secret. But Hirasawa says he thinks the present system of collecting vast amounts of data for empirical prediction is "basically correct". Geller expects the report to preserve the status quo. "No doubt the seventh five-year plan will be the same as the sixth five-year plan, which was the same as the preceding five plans", he says.

The Japanese situation stands in stark contrast to that in the United States, where hundreds of scientists, engineers, and government planners participated in a recent review of the US Geological Survey (USGS) earthquake hazard reduction programme. "Ten to fifteen years ago, US researchers put a lot of effort into an empirical approach to earthquake prediction just like Japan", says Hiroo Kanamori, director of the seismological laboratory at the California Institute of Technology in Pasadena. But current emphasis is on understanding the basic mechanisms of earthquakes and on research related to disaster mitigation.

In an effort to open the process slightly, the draft review of Japan's programme is being sent for the first time to six 'outside' reviewers, including the heads of the seismological and volcanological societies of Japan, and Masuo Iida, a former MESC official, who has been very critical of the present programme. However, the reviewers will have only a few weeks to comment on the draft report before it is submitted to the government in July, and substantial changes cannot be expected. And the people who chose these reviewers are the same people who set up the drafting committee.

David Swinbanks

Power company leaves environmental legacy

London

ENVIRONMENTAL researchers across Britain have received some £1.8 million (US\$3.1 million) worth of equipment as a result of the privatization last year of Britain's public utility. But their happiness with the windfall is tempered by the dissolution of one of the world's largest acid rain research laboratories, and the government's failure to compensate for such losses by increasing its spending on environmental research.

Last year, National Power — formerly part of the publicly owned Central Electricity Generating Board (CEGB) — decided to scale down dramatically its research activities after privatization (*Nature* 352, 460; 1991). Research programmes that were meant to solve national problems or deal with international issues fell outside the company's concern, it announced, and were to be discontinued. Only research essential for its core

business of generating electricity would remain.

Since then, a large proportion of the equipment has been distributed through the Natural Environment Research Council (NERC). Around 160 items — including a number of large, sophisticated controlled environments for studying the effects of pollutants on plants and their movements through soils — have been shared by 27 different groups, including university departments, the Forestry Commission and NERC's own terrestrial and fresh-water institutes.

In a separate deal, staff and £30,000 in equipment have gone to Imperial College, London, to form an atmospheric chemistry research unit. The package includes a share in a Jetstream aircraft, fitted out for observing plumes from industrial sources or cities and for large-scale studies of tropospheric pollutants.

The group, which also brought with it

£50,000 in start-up money from National Power, is keen to continue the work it was doing under the company. This includes a number of cooperative ventures with the European Commission, such as examining the dispersion of pollutants over complex terrains such as the Alps.

National Power also put money into a plan by senior staff at its Fawley marine laboratories to transform themselves into a commercial operation, offering contract research and monitoring services to power and water companies both in the UK and in Europe. Donations of more conventional laboratory equipment went to university departments and hospitals in Britain and to Charles University in Prague.

Although this seems like a story where everyone benefits, there are some discordant voices. Some observers say that National Power is just off-loading equipment otherwise intended for the scrapheap. They add that the benefit does not outweigh the loss of the research work carried out under the old CEGB. Certainly one of the world's largest acid rain teams (with an annual budget of £4.5 million) has been lost, and a vague statement by the government that it would take up the slack in funding environmental research dropped by the power companies has not been realized.

But officials at National Power say that much of the CEGB research had come to a natural end. Work related to decision-making — the consequences of continuing with large, coal-fired plants and adopting combined-cycle gas turbines — ended because the decisions had been made. Similarly, clear environmental objectives from the EC removed the need for the company to set its own standards. What still needs to be done, they say, can be farmed out.

"Increasingly, universities and other institutions are more geared up to provide that research, so it is easier to contract out," says a spokeswoman for the company. National Power's remaining scientists, relocated to its Swindon offices, will be involved in "keeping a finger on the pulse" of what is going on but will do very little hands-on research. The utility company has not decided how much it intends to spend on outside research. **Ian Mundell**

Correction

In the News story "NIH staffer finds warm reception on other side" (*Nature* 355, 98; 9 January 1992), the date mentioned for the meeting between National Institutes of Health (NIH) officials and Congress should have been 4 November 1991. *Nature* was similarly wrongly informed that there had been an "exchange" of letters between the NIH and Congress; in fact, NIH have yet to respond to the congressional letter. Elsewhere in the story, an ambiguous sentence could have been read to mean that Walter Stewart, a NIH employee, was not paid by NIH while he was temporarily assigned to work for Congress. In fact, Stewart remained on the NIH payroll throughout his assignment.

SCIENTIFIC MISCONDUCT

New court challenge for OSI

Washington

THE embattled scientific misconduct office at the US National Institutes of Health (NIH) has been sued again, this time in a class-action filing that could affect misconduct cases throughout the United States.

Herbert Needleman, a University of Pittsburgh lead researcher, filed suit late last month against his university and NIH, claiming that the NIH misconduct rules are being imposed retroactively and in violation of federal rule-making regulations. Needleman is under investigation by the university at the request of NIH's Office of Scientific Integrity (OSI). The university is reviewing allegations that Needleman misrepresented data in two studies — one in 1979, the other in 1990 — on the effect of lead on children's intelligence. A victory for Needleman would invalidate the OSI rules nationwide.

By filing a class-action suit, Needleman and his lawyer, James Lieber of the Pittsburgh law firm Lieber & Hammer, are claiming to represent "all those scientists who are or will be charged with misconduct" by OSI, Lieber says. Class-action suits are permitted in cases where the number of parties claiming substantially similar injury are too numerous to name. In practice, that has traditionally meant about 50 or more.

Needleman wants to have the NIH rules invalidated and his investigation halted, but he is not seeking monetary damages. One advantage to plaintiffs in a class-action suit is that legal costs can be shared, but such a savings is not the motivation in

this suit, says Lieber. "It's not an economic question, it's a legal question", he says.

OSI's regulations have been declared illegal in one previous case — that of James Abbs, a University of Wisconsin researcher — but in that case (which is still under appeal) OSI was prevented from exercising its authority only in one district in Wisconsin. A similar suit, if filed in a federal court in Washington DC, would have nationwide impact.

But Needleman wants to halt the university's investigation as well, so he is filing in a Pittsburgh court. Calling a class action suit an "altruistic" action, Lieber says the blanket approach is intended to save other researchers from filing "piecemeal litigations up and down the country."

Needleman is claiming that a definition of misconduct published in 1990 by OSI should not be applied to research that was conducted a decade earlier. The rules are also impossibly vague, Lieber argues; catch-all definitions such as "any other practices that seriously deviate from the norm" are invalid under a US law that governs such ambiguous phrases, he says.

Responding to faculty support for Needleman, Pittsburgh agreed last month to open the investigative proceedings to the general public. Although open hearings (which begin on 13 April) do not satisfy Needleman's demands for due process, he hopes that either his suit or continuing negotiations with the university will eventually result in such rights as cross examination and written charges.

Christopher Anderson

US Senate votes to overturn research ban

Washington

THE end of three-year-old US ban on federal research using tissue from human fetuses is in sight after the US Senate voted overwhelmingly last week to overturn it.

President George Bush has threatened to veto similar legislation in the past, arguing that fetal tissue research encourages abortion. But the margin of victory in the Senate — 87 to 10 — suggests that the Congress can obtain the two-thirds majority necessary to override a presidential veto. Supporters of fetal tissue research met with White House aides last week to urge Bush not to veto the bill, which will be reconciled with a similar version passed last year by the House of Representatives and then sent to the president.

The vote in the House last July was 274 to 144. The Senate vote suggests that advocates of fetal tissue research have made considerable progress in the past nine months.

A heavy lobbying campaign for the bill, which is part of a larger measure to reauthorize parts of the US National Institutes of Health (NIH), was led by the

American Federation for Clinical Research (*Nature* 355, 189; 1992). Supporters were jubilant at the victory. "This represents an excellent example of scientists from academia educating legislators to separate research from politics," says Andrew Hoffman, the federation's president and a Stanford University professor.

The bill also dealt with the controversial question of federal funding for studies on the sexual behaviour of certain populations. An amendment proposed by Senator Jesse Helms (Republican, North Carolina) was adopted that prohibits funding for two surveys — one involving teenagers and one that would sample the US adult population. The surveys became political footballs and were suspended last year after legislators protested that they represented an invasion of privacy. However, a second amendment from Senator Paul Simon (Democrat, Illinois) was also passed that would make such surveys more difficult to kill once they had undergone scientific peer review.

The combination of Simon's amendment and the "research freedom" language

of the fetal tissue portion of the bill will set up an approval mechanism for sex surveys that should limit future political meddling, predicts Howard Silver, president of the Consortium of Social Science Associations. "The existing proposals, as we know them, are dead," he says. "But they can be reintroduced and should do much better under the new mechanisms."

On the eve of the vote, Senator Mark Hatfield (Republican, Oregon) took aim at both the biotechnology industry and NIH with a proposal to prohibit patents on any part of the human body and to impose a three-year moratorium on patenting genetically engineered animals and organisms while a new biomedical ethics board evaluated the implications of such patents. Beyond ethical and safety concerns about genetically engineered organisms, Hatfield questioned NIH's recent patent filing for thousands of human cDNA sequences. After being lobbied intensely by industry and the Bush administration, Hatfield agreed to introduce and then withdraw the amendment in exchange for a future hearing on the gene patent issue by the patent subcommittee of the Senate Labor and Human Resources Committee. The bill now also asks the congressional Office of Technology Assessment to prepare a report on the subject.

Christopher Anderson

BRAZIL

Cabinet reshuffled again

São Paulo

THE president of Brazil, Fernando Collor de Mello, has appointed a new secretary of science and technology as part of the most recent upheaval within his Cabinet. The move was greeted positively by the scientific community, which sees the appointment as a promise of increased support.

The new secretary is Hélio Jaguaribe, a sociologist and member of an opposition party. He succeeds Edson Machado, a career bureaucrat with no political ties. Collor's wish to accommodate Jaguaribe's party is expected to extend beyond a seat in the cabinet to providing additional money for science.

Jaguaribe has written a book, *Brazil: Reform or Chaos*, about the threat to the country from the enormous disparity between rich and poor. Some members of his party, the social democrat PBDB, have refused to cooperate with Collor, but Jaguaribe has discussed this subject and other issues with him.

The director of the Latin American Federation of Physics Societies, Gil da Costa Marques, praises his "independence". But the consensus view is one of cautious optimism. "I hope it means change", says Paulo Milton Barbosa Landim, the rector of São Paulo State University. Referring to his displeasure at the current level of government support for research, Landim added, "the last secretary did not do a good job."

The reshuffling, which so far has involved eight Cabinet ministers, further strengthens the hand of Machado's predecessor as science secretary, physicist José Goldemberg. Goldemberg was promoted last year to Minister of Education, and last month took on the additional job as environmental secretary after the dismissal of José Lutzenberger (see *Nature* 356, 278; 1992). A last-minute appointment after Collor took office in March 1990, Goldemberg has since become a member of the president's inner circle of advisors.

Goldemberg was in New York last week when the latest changes were announced. He was looking for money for environmental projects in his country, something that Lutzenberger was unwilling to do. In fact, Lutzenberger was fired for telling the industrial world not to lend money to Brazil for conservation projects because it might end up in the hands of corrupt politicians. He said that problems existed even within his own agency for environmental protection, the Brazilian Institute of the Environment and Natural Renewable Resources.

Although Lutzenberger refused to provide names, the charges appear to have substance. Last week, for example, the institute's office in Rio de Janeiro was asked to explain how a 'farm' it had purchased to extend a natural preserve is actually located under the Atlantic Ocean.

Ricardo Bonalume

UK HIGHER EDUCATION

Faraday, Newton... and Parnaby?

THE name of Parnaby has joined those of Newton and Faraday as titles for centres of excellence designed to revitalize the training of graduate scientists and engineers in the UK.

The Parnaby centres would play host to the revamped engineering doctorates announced last week by the Science and Engineering Research Council (SERC). Named after John Parnaby of Lucas Applied Technology, who headed the SERC working party set up to revise the UK's engineering doctorates, they are seen as having a role parallel to the Faraday Centres and Newton Institutes that have been proposed by various political bodies (*Nature* 355, 757; 1992). All are loosely based on the German Fraunhofer Institutes.

The current system of training engineers in Britain is seen as not meeting the needs of the manufacturing industry. The new doctorates will last four years instead of the customary three, and will combine close industrial cooperation, courses in a range of subjects and training in management techniques. The SERC will support three pilot studies, each with 10 students, to test the idea.

Ian Mundell

A steady course towards the next tokamak

Tokyo

RESEARCHERS at the Japan Atomic Energy Research Institute (JAERI) north of Tokyo have begun to design a steady-state tokamak to succeed the institute's aging JT-60 tokamak. Proponents argue that it is needed to preserve Japan's role in developing fusion devices for the next century, but critics question whether it is the best use of scarce resources.

Fuelled with money appropriated for the fiscal year that began last week, JAERI researchers will join their industrial counterparts at Hitachi, Toshiba and Mitsubishi in completing the design of a steady-state tokamak that is intended to be a stepping stone between JT-60 and the International Thermonuclear Experimental Reactor (ITER). The \$6,000-million ITER project is expected to be a combined effort of the United States, Japan, Europe and the Soviet Union.

The aim is to make a device that can maintain a heated deuterium-deuterium plasma for 100 to 1,000 seconds, much longer than the short bursts of a few seconds that current tokamaks like JT-60 achieve. And it would be a final step before making a device such as ITER that can exceed the 'break-even' point and produce more energy than it consumes.

The JAERI initiative, which has yet to win government approval, matches a similar move to build a steady-state device in the United States (see *Nature* 356, 96; 1992). It is motivated partly by fear that the institute's scientists will be left without an advanced fusion device of their own when JT-60 reaches the end of its life.

Just as the US proposal for a Tokamak Physics Experiment came after fusion scientists failed to win funding for the more ambitious and more costly Burning Plasma

Experiment, the JAERI steady-state plan comes after the institute's plan to build a Fusion Experimental Reactor was shelved by Japan's Atomic Energy Commission. The commission instead cast its lot with an effort that would lead directly to the ITER prototype reactor.

US scientists are driven by the knowledge that the country's leading



A worker stands inside the JT-60 machine during its recently completed renovation.

tokamak, in Princeton, will reach the end of its scientific life in 1995. But the Japanese face a less urgent situation. The JT-60, which has just undergone a major renovation, is thought to be good for another dozen or so years.

Although the planned steady-state device will use the existing power facilities

for the JT-60, most of it will be built from scratch, including the vacuum vessel and magnets. Unlike the current tokamaks, the new device will use superconducting magnets and special robot equipment. The superconducting magnets generate stronger fields in a more compact area and at lower operating costs, and the robots are the only way to maintain the walls of the tokamak, which become radioactive after confining the deuterium-deuterium plasma for long periods of time.

JAERI researchers declined to comment on the expected cost of the new machine. But US scientists who saw preliminary design plans of the Japanese tokamak at a meeting last November estimate that the Japanese machine would cost about \$1,000 million, or double the price of the US machine. The US Burning Plasma Experiment would have cost \$1,500 million.

The Japanese machine is expected to take about four years to build. JAERI officials would like to be using it before the end of the decade, meaning that a formal proposal must be submitted to the Japanese government in the next few years if that timetable is to be met.

Despite all the planning that has taken place, some Japanese fusion scientists question whether this project should move forward. Atsuo Iiyoshi, director of Japan's new National Institute for Fusion Science, says "I don't disagree with their plan". But if fusion budgets are going to be limited, he says, it would be better "to divide roles" among existing programmes.

One such effort is a giant helical fusion device under construction at Iiyoshi's institute. Its prime purpose is to examine steady-state experiments in fusion (see below). Iiyoshi suggests that JAERI could conduct research involving burning plasmas by using tritium in the JT-60, while Iiyoshi's new fusion institute looks at the characteristics of steady-state experiments.

But JAERI researchers do not want to burn tritium in the JT-60. That is one of the research aims of scientists at Princeton and the JET (Joint European Torus) tokamak in Britain, they say.

The fate of the JAERI plan is uncertain. The timing of a proposal for a follow-up to JT-60 is "delicate", says Ikuo Makino, director of the Technology Development Division of the Atomic Energy Bureau at the Science and Technology Agency, which is in charge of JAERI. Such a project, he says, must not compete with the construction of ITER, which could begin in the latter half of the 1990s for completion by 2005. Still, it is clear that JAERI has made the steady-state machine a priority.

David Swinbanks

Helical fusion machine on course

JAPAN's plans to build the world's largest helical fusion device are proceeding "almost on schedule", says Ikuo Iiyoshi, director of Japan's new National Institute for Fusion Science. The device, which is under construction at Toki in Gifu Prefecture near Nagoya (see *Nature* 335, 104; 1988), represents a key part of Japan's fusion research that is funded by the Ministry of Education, Science and Culture. In contrast, research at the Japan Atomic Energy Research Institute (JAERI) is funded by the Science and Technology Agency.

The new institute will have a staff of 280, including 150 researchers, when it opens in 1995. It is based temporarily at Nagoya University. The device will cost ¥50,000 million (US\$370 million), but only about one-tenth of that amount has been spent so far.

This year, some ¥4,500 million (US\$33 million) will be spent on the fabrication of machines to make superconducting magnets, and a similar amount for new buildings. The institute also plans to buy a supercomputer, Iiyoshi says.

The helical device will produce a magnetic energy of 2 gigajoules and will be by far the largest such device in the world. Its purpose is to examine the physics of plasmas under steady-state conditions, a mission not linked directly to the development of commercial fusion reactors.

D.S.

Report backs whistleblower

Washington

THE US Department of Energy (DOE) has sided with a metallurgist at one of its research laboratories who claims that he was hounded out of his job at Argonne National Laboratory after repeatedly raising questions about faulty and misleading data on an experimental nuclear reactor programme.

In a report made public on 2 April, the DOE's Office of Nuclear Safety (ONS) sided with James A. Smith, who says that fundamental errors were being made and published in work on a breeder reactor programme known as the Integral Fast Reactor (IFR). Argonne, which is 45 miles west of Chicago, is operated for DOE by the University of Chicago.

The 122-page ONS report also agreed with Smith's claim that he was threatened with the loss of his job after writing a memorandum suggesting that laboratory procedures to detect errors in data need to be improved. Smith resigned his position as an experimenter in August 1990 after being suspended. He admits that he is not interested in returning to the Argonne reactor, which is located at the Idaho National Engineering Laboratory.

Argonne strongly disputed the report from the ONS and says that Smith was discharged for not doing his job. Officials say that Smith became obsessed with the desire to refute every aspect of a fellow scientist's work, leaving him with little time to perform his own duties.

Smith's actual scientific concerns, which related to the melting points of the elements used in reactor fuel assemblies, seemed secondary to those that related to the scientific atmosphere of the laboratory. Smith said several of his colleagues told him he should stop worrying about the quality of work at the laboratory and consider his job as a way to earn a decent salary and live in a nice location.

But Smith rejected their advice. "I didn't get a Ph.D. in metallurgy to go into some high-tech yuppie welfare programme," he says.

The DOE said that it did not find any evidence of fraudulent intent on the part of Argonne personnel, but that some work had been published or disseminated that was misleading. It said that the laboratory has a policy not to retract, correct or qualify research even when errors or serious questions about its validity are uncovered. The ONS report also found that Smith described accurately a scientific culture at the laboratory that was "oversensitive to internal politics, real or imagined slights and other social considerations." That culture, it says, undermines peer review and adherence to scientific accuracy.

A spokesman for the laboratory denies emphatically that Argonne has any such

policy with respect to refereed journals but that the material in question appeared in an in-house newsletter. "It's the written equivalent of walking down the hall and talking to your colleague", says the spokesman. Retractions or corrections are made at the discretion of the author and the editors of the newsletter, he added.

Argonne also denies that it intentionally published any misleading material, and asserts that it stands solidly behind peer review and adherence to scientific principles and accuracy. In its defense, the laboratory pointed to an outside evaluation of the technical issues that Smith raised.

The technical panel, convened after the ONS completed its report last December, concluded that none of the concerns that Smith raised threatened the safety of the laboratory principally because IFR, a metal-fuelled, liquid sodium-cooled 'fast' reactor, is at least 20 years away from production. But the DOE report said that the kinds of errors and practices that Smith identified could be repeated in other programmes at the laboratory.

Smith also questioned the independence of the technical evaluation committee. He pointed out that two of the panel's three members also sit on a special advisory panel to the reactor project that reports to the University of Chicago's board of governors. He says that the committee ignored the "gross blunders" that he raised.

"It's a case of the foxes guarding the chicken coop," Smith says of the technical panel review. "I'm shocked that the head of a national laboratory would even accept such a report, let alone tell people it vindicates his laboratory."

A spokesman for Argonne concedes that such a conflict-of-interest might exist "if someone had a vested interest in doing so". But he defends their presence on the evaluation committee by saying that it amounts to only a tiny amount of their professional duties and that the pool of experts in the field is small. It would not be prudent of them to avoid potential or real problems, he adds.

In spite of its disagreement with DOE on Smith's case, Argonne has agreed to implement the recommendations in the ONS report. They include developing an explicit policy of correcting known errors in its publications and issuing appropriate caveats when questions about its research are raised. The laboratory was also ordered to train its personnel to follow accepted principles of peer review that include how to resolve disputes between scientists. Employees will also be taught to recognize and respect those activities in which a whistleblower can be engaged without fear of retaliation.

David Kramer

China drives forward

Beijing

ONCE China was a country that praised manual labour above brain work. Not any more. Last month the city government of Zhuhai in South China's Guangdong Province awarded an Audi automobile, an apartment and hundreds of thousands of Yuan (5.5 to the US dollar) to each one of a group of scientists for their outstanding contributions to their country's economic development.

Chinese intellectuals have complained for years about their low pay, poor living conditions and lack of social respect. But the program in Zhuhai, which is expected to continue indefinitely, marks an end to that situation. With its emergence as a global trading partner, China is looking for ways to encourage researchers to turn their ideas into products.

Putting a price tag on scientific progress is not easy, of course. The Zhuhai awards have raised questions about the value of those who engage in basic research. Should those whose work leads to commercial products be paid more than those whose achievements are mainly intellectual? Such ideas may seem far-fetched in a country that once put equality above all other civic principles, but increasingly they reflect what government officials are trying to accomplish.

You Qin Li

Australia steps up

Sydney

AUSTRALIAN engineers and scientists who work at large companies will receive salary increases of as much as 35 per cent this year under a ruling by the government commission that sets wages for a large portion of the country's labour force. The decision by the Industrial Relations Commission is viewed as a fundamental realignment of the pay rates for technical workers, in keeping with the government's desire to strengthen the country's research capacity and make science more attractive to students and others entering the labour force.

Under the new rates, the minimum starting salary for an engineer will rise by 35 per cent, to US\$21,470, and the pay of an engineer with four years' experience will increase by 29 per cent, to \$US26,430. The increase will affect about 25 per cent of the country's estimated 25,000 professionals whose employers are subject to the commission's ruling. Scientists and engineers at small companies negotiate their own salaries, which are frequently higher than the levels set by the commission.

Last year, academics throughout the country received a one-time salary increase of 16 per cent, and in 1990 government scientists received a similar increase. Both increases are supplemented by cost-of-living rises tied to the annual rate of inflation, now running at 4 per cent. Mark Lawson

Human rights in Cuba

SIR — C. J. Brown and I have compiled a directory of 63 cases of Cuban political dissidents abused by psychiatric methods¹. Thirty-one of these cases could be documented from the publications of Amnesty International and Americas Watch, and revealed a systematic pattern of abuse and torture. Eleven dissidents were forcibly subjected to electroconvulsive therapy, 16 were forcibly given high doses of psychotropic drugs, most chlorpromazine, and four were interned with criminally insane patients in the Havana Psychiatric Hospital, and there threatened with electroshocks to frighten them and secure confessions of political offences.

The political nature of the dissidents' offences cannot be doubted. Eight were interned after unsuccessful attempts to leave Cuba illegally; another ten were interned for unspecified counterrevolutionary activities; eight were accused of writing anti-Castro graffiti and distributing human rights leaflets; and the rest were accused of various offences, including contempt for the regime, refusal to serve in the Cuban armed forces and alleged participation in plots to kill Fidel Castro. The age of the dissidents ranged from 16 to 61. A variety of occupations and professions were represented but two groups stood out: students and journalists (5) and writers (6).

Most of the abuses took place in the forensic wards of the Havana Psychiatric Hospital, the Gustavo Machín Hospital in Santiago de Cuba and the Combinado del Este Prison Hospital in a suburb of Havana. The forensic wards of these hospitals are under the direct control of the Cuban Department of State Security, the Cuban KGB. The duration of confinement varied from three days to five years in the cases of two dissidents. Cuban penal law allows for placing prisoners for up to a month at the Havana Psychiatric Hospital for psychiatric evaluation. But in more than half the cases analysed, the period of internment exceeded a month, in four cases was longer than a year. Only 11 of the cases investigated were diagnosed by Cuban authorities as suffering from mental illnesses. Three dissidents never received a psychiatric evaluation, and four others were diagnosed as sane but were nevertheless sent back to the psychiatric wards for further confinement. One dissident was diagnosed as "apathetic to socialism" and another as suffering from "delusions of being a defender of human rights".

None of the Cuban dissidents was diagnosed as suffering from sluggish schizophrenia, the diagnosis that was often used in the Soviet Union to intern

dissidents in psychiatric hospitals².

Although electroconvulsive therapy is appropriate for some mental illnesses, such as severe depression and catatonic schizophrenia, this therapy is systematically applied in Cuba to punish past dissident actions. Electroshocks are applied without medical supervision and under barbaric conditions. The dissidents were not given muscle relaxants to prevent fractures and, while forcibly held to the floor by criminal orderlies, their bodies were doused with water to increase electrical conductivity and electroshocks administered on floors covered with vomit, urine and excrement.

In the case of the 61-year-old businessman Eugenio de Sosa, who received 14 electroshock sessions as punishment for smuggling messages from prison, most of the electroshocks were applied to his testicles. Another dissident, the truck driver José Morales, received one of his 14 electroshocks while submerged in a water tank used for cattle. Jesús Leyva, a union leader, received 24 electroshock sessions. In an interview broadcast on television in Florida, Leyva claimed that after one electroshock session he did not recognize his wife or his relatives. Roberto Bahamonde, an agricultural engineer, underwent eight electroshock sessions after writing Fidel Castro a private letter suggesting that material incentives be instituted to increase production. He told journalist George Gedda of Associated Press that it took him years to recover his memory.

Equally disturbing were the large doses of psychotropic drugs that dissidents are forced to ingest. A major problem with these drugs is their side effects. One such side effect interferes with the extrapyramidal functions of the brain responsible for movement, locomotion and coordination. The danger of becoming a human vegetable by experiencing a chemical lobotomy from long-term exposure to drugs is real³. Several of the Soviet victims of psychiatric abuse exhibited these side effects³, as did one of our cases.

The time has come to insist that these practices be stopped by the Cuban government. Political use of psychiatry in Cuba must end.

ARMANDO J. LAGO

*Of Human Rights,
Box 2160, Hoya Station,
Georgetown University,
Washington, DC 20057, USA*

1. Brown, C. J. & Lago, A. M. *The Politics of Psychiatry in Revolutionary Cuba*. (Transaction Books, New Brunswick, New Jersey, 1991).
2. Firestone, H. *Soviet Psychoprisons* (Norton, New York, 1979).
3. Bloch, S. & Reddaway, P. *Psychiatric Terror* (Basic Books, New York, 1977).

In situ situation

SIR — In a recent issue of *Nature* (355, 163; 1992), we were attracted by the title of a paper, "Rapid-time-course miniature and evoked excitatory currents at cerebellar synapses *in situ*", especially as we were looking forward to welcoming new members to the otherwise ever-shrinking club of people performing intracellular recordings in whole animals. But our surprise and disappointment were great when we read the paper and discovered that the work had been done *in vitro*, in a slice. We do not have any qualms about the quality of this work, but we do feel that the use of the term '*in situ*' was clearly misleading in this case. Indeed, it is difficult to see why it was added to the title.

We should like to suggest that '*in situ*' should not be used to refer to work done *in vitro*, as there are already precedents in the literature — for example: Do *et al.* *Neuroscience* 8, 2226 (1986) Herreras, J. *Neurophysiol.* 64, 1429 (1990); Herrling *et al.* *Neuroscience* 31, 213 (1989); Hernandez *et al.* *Expl Brain Res.* 85, 66 (1991) — where this term is used to refer to work carried out in whole animals. We suggest following the practice of the *Journal of Physiology* where work carried out in brain slices or cultures is referred to as '*in vitro*'. Work done on whole organisms should be termed '*in vivo*' or '*in situ*'.

PAUL L. HERRLING

*c/o Sandoz Research
Institute Berne Ltd,
CH-3001 Berne,
Switzerland*

T. E. SALT

*Department of Visual Science,
Institute of Ophthalmology,
Judd Street, London WC1H 9QS, UK*

Speaking volumes

SIR — VCH, the publishers of the *European Journal of Immunology*, point out that, contrary to my statement (*Nature* 354, 346; 1991), their journal cover does indeed display the volume number. I acknowledge my error, but the volume number is in small type at the foot of the page among a plethora of other data, whereas the issue number is presented at the head of the page, in the same large typeface as the journal title and year.

The fact that I, an experienced professional librarian, can make such a mistake only goes to strengthen the case for some sort of standardization of bibliographic data.

SHEILA M. MOULD

*MRC Human Genetics Unit,
Western General Hospital,
Crewe Road,
Edinburgh, EH4 2XU, UK*

Forensic use of PCR in Italy

SIR — Following the admission of the use of evidence obtained through the polymerase chain reaction (PCR) in British courts¹, attention has been drawn to the forensic use of PCR in Italy. Dallapiccola *et al.*² reported that at the Balsorano-Avezzano court case they were able irrefutably to involve a man in a murder by a PCR-driven DNA analysis, providing arguments in favour of PCR use and a reliable protocol for routine work.

Having acted as expert witnesses at that trial, we believe that the evidence produced was of dubious value and the protocol deserved severe criticisms.

(1) Although they recommended great care in producing forensic evidence by DNA profiles, Dallapiccola *et al.* carried out a technically flawed analysis. Overlooking elementary rules to assert a match between forensic samples, they type amplified VNTRs (variable number of tandem repeats; for example D1S80) and hypervariable restriction fragment length polymorphisms (D2S44) — endowed with quasi-continuous variability — on small-size electrophoresis with inadequate resolution, rough molecular mass ladders and no allelic standards. Having no population database for their markets, they could not draw any warranted probability estimate. Being ignorant of the precautions to adopt in referring statistical inferences to substructured populations^{3,4}, they overlooked the fact that the Appennine mountain community where the crime was committed (Borgo Case Castella di Balsorano, 97 inhabitants in 1990) is in fact a genetic isolate with a high rate of inbreeding. Against this background, they issued an arbitrary probability figure (1.6×10^{-4}), which nonetheless influenced the jury. In short, they fell into just the situation of 'worthless evidence' and 'unreliable conclusions' against which they warn.

(2) We are very surprised to find that in their letter² the impact of DNA analysis in this court case is even more exaggerated than in their original, written report at the trial, as they now quote hyperbolic, seven orders of magnitude higher odds ($1:1.7 \times 10^{-11}$) to indict the defendant.

(3) Last but not least, the protocol they used is not suitable for forensic analysis. Largely based on scarcely polymorphic systems, not validated for forensic purposes (APO-CII, 3'CACA, distrophin, VNDR18 and others), it lacks standardization, quality controls and mutual objectivity. By disregarding so many obligatory issues to produce scientific evidence in court⁵, Dallapiccola *et al.* rendered the wrong service to a good cause (forensic use of PCR). As part of a

group promoting standardization and quality assurance in the field (the European DNA profiling group, EDNAP) we denounce this example of redundant, as well as unsuitable forensic analysis.

ANGELO FIORI

VINCENZO L. PASCALI

*Department of Forensic Medicine,
Catholic University of Sacred Heart,
largo F. Vito, 1,
00168 Roma, Italy*

1. Hagelberg, E., Gray, I. C. & Jeffreys, A. J. *Nature* **352**, 427–429 (1991).
2. Dallapiccola, B., Novelli, G. & Spinella, A. *Nature* **354**, 179 (1991).
3. Lewontin, R. C. & Hartl, D. L. *Science* **254**, 1745–1750 (1991).
4. Brookfield, J. *Nature* **355**, 207–208 (1992).
5. *Nature* **350**, 95 (1991).

Air pollution and cancer risk

SIR — It is pleasing to learn from Paolo Paoletti *et al.* (*Nature* **355**, 290; 1992) that serious efforts are under way within at least one country of the European Communities (Italy) to determine the effects of air pollution upon health, including cancer risk.

The British view is that the cancer risk to the general population from vehicle-derived air pollution is so small as to be unmeasurable (F. Godlee *Br. med. J.* **303**, 1539–1543; 1991). Lung cancer in Britain is now attributed almost solely to active and passive smoking (with, perhaps, a small contribution from domestic radon). The million-odd tonnes of known or probable carcinogen churned out into the air by road transport annually is deemed innocuous.

Yet observations that seem to conflict with this view (urban are higher than rural lung cancer rates, risk is highest in the most densely trafficked urban areas and so on) are dismissed on the basis of local differences in socioeconomic factors or smoking behaviour. But it is difficult to explain why lung cancer rates in rural China are an order of magnitude lower than in Britain, and not much different between smokers and nonsmokers¹.

Interestingly, when the increase in lung cancer rate in developed countries was first noticed earlier this century, it was a moot point whether air pollution (particularly from the new-fangled diesel lorries and buses) or tobacco smoking was to blame. That the entire blame has been laid at the door of smoking alone has been a source of mystery to medical sociologists². The likely possibility of synergy between powerful mutagenic factors adsorbed on diesel particulates and cancer-promoting factors in cigarette smoke has never been closely assessed

and, over the decades, any possible contribution of air pollution to lung cancer has been carefully edited out, perhaps, I have suggested³, to the quiet but forceful involvement of groups outside medical science with a natural interest in vehicular air pollution proving to be harmless.

Wide-ranging prospective surveys such as those proposed by Paoletti and his colleagues may indeed answer the question: "Does air pollution cause cancer?", but we also need cross-sectional surveys of risk factors in patients. As air pollution caused by traffic tends not to be homogeneous, it should be feasible not only to ask lung cancer patients "How many cigarettes do you smoke each day?" but to ask local traffic engineers how many vehicles pass their residence.

SIMON P. WOLFF

*Department of Clinical Pharmacology,
UCMSM,
University College London,
5 University Street
(The Rayne Institute),
London WC1E 6JJ, UK*

1. Katz, R. & Chunxiang, Z. *Hlth. Phys.* **51**, 457–468 (1986).
2. Vandenbroucke, J. P. *Am. J. Epidemiol.* **133**, 426–427 (1991).
3. Wolff, S. P. *Am. J. Epidemiol.* (in the press).

Authors and egos

SIR — Christopher Anderson's report (*Nature* **355**, 101; 1992) on publication frequency by scientists should be cause for grave concern in the scientific community. Authorship of a scientific publication is not a reward for having assisted in some way, however trivial, in making a research report possible.

If one is conducting field research in the Andes, for example, the muleteer hired to provide access to the study area obviously makes a vital contribution to the research effort but few would argue that this effort warrants co-authorship of any publications resulting from the expedition's scientists. Similarly, someone who is awarded a grant does not automatically merit co-authorship merely because the funds made possible the research of others, any more than the largesse of an individual private donor who provides funds for research would merit inclusion on an author line.

If this were not a serious problem in science, the almost childlike attempts to feed enormous egos would be silly, indeed. Does anyone really believe that someone authors a paper every few days? Clearly not. Perhaps it is time for reviewers of manuscripts and grants to stop playing this little game and deduct points for patent ego engorgement.

MICHAEL A. MARES

*Oklahoma Museum of Natural History
and Department of Zoology,
University of Oklahoma,
Norman, Oklahoma 73019, USA*

Global environmental engineering

Ralph J. Cicerone, Scott Elliott and Richard P. Turco

All the signs are that global ozone depletion is increasing. Ideas to mitigate the problem that at first glance may seem far-fetched deserve more serious consideration and a scientific process of evaluation.

THERE is new evidence of a serious problem in the stratospheric ozone layer. Direct and immediate biological damage due to increased ultraviolet light in Antarctic waters has now been reported¹ and ozone in Arctic and northern mid-latitude regions may be threatened — even as international agreements (the Montreal Protocol), national plans and those of industries are being implemented and accelerated to stop the flow of chlorofluorocarbons into the atmosphere. It may become necessary to develop more aggressive strategies if damage from reduced ozone becomes intolerable. Should these include gross manipulation of the environment?

Many ideas about global environmental engineering have been casual, scientifically unsound, dangerous or all of these — for example suggestions such as exporting ozone to the stratosphere or 'sucking' chlorofluorocarbons out of the atmosphere. But serious suggestions are beginning to be made. Some interest arises from simple faith in technological fixes, some from technologists who wish to help, and some from large corporations with significant scientific and engineering capabilities in search of business opportunities. There are several major companies investigating possible projects, at least one of which has been granted a patent². Ideas are also being raised by people concerned about environmental degradation.

By all measures, the Antarctic ozone hole is becoming a more serious problem. Average amounts of ozone measured each October have decreased sharply during the past 15 years; daily maximum and daily minimum amounts have also decreased; and the duration of these decreases has increased. Further, the area of reduced ozone has spread well into southern midlatitudes, possibly because of the export of ozone-poor air from the Antarctic stratosphere late each austral spring. Episodes of reduced stratospheric ozone over Australia have been recorded which suggest an Antarctic origin. Increased intensities of ultraviolet radiation have been measured beneath the hole, with intensities more typical of the summer occurring in early spring. Meanwhile, compelling experimental and theoretical evidence indicates that chlorine from chlorofluorocarbons is the main ozone-destroying agent and that the activity of chlorine is

maximized by chemical reactions on the surfaces of polar stratospheric cloud particles. The fact that chlorofluorocarbons survive for about 100 years in the atmosphere implies that stratospheric chlorine will not return to pre-ozone hole values until late next century. Therefore, the Antarctic ozone hole will probably continue to develop each year, and the amount of ozone in the Southern Hemisphere will continue to decrease.

We are investigating a concept to close the ozone hole — the idea of injecting thousands of tons of alkanes (such as ethane or propane) into the polar stratosphere over Antarctica for about a month each year³. Alkanes (and their decomposition products) are relatively short-lived, so would not spread globally. There are many uncertainties underlying this idea, for example, information about the formation rates of polar stratospheric clouds and of chemical reactions on the cloud surfaces are badly needed. But we have attempted to evaluate key processes, to look for sensitivity to various assumptions and to look for possible side effects by means of a computer model. Our model shows that although ozone depletion could be prevented or slowed through annual alkane injections, there are other plausible outcomes in which small injections could worsen ozone depletion. It has now been proposed that there is a direct reaction of HCl and HOCl on the surfaces of polar stratospheric clouds⁴: this reaction could also alter the response to alkanes. So although our model is simplified, and serious questions of feasibility are not addressed, we suggest nonetheless that exercises like this are essential first steps in the responsible exploration of the feasibility of environmental engineering.

For an effective debate, publication of ideas in peer-reviewed journals and peer-review of research plans is essential. Detailed feasibility studies should wait until the underlying scientific concept is documented as well as reviewed critically and repeatedly. Analyses of engineering costs or of timetables for action are premature at this stage. Ideas for global environmental engineering are likely to be badly flawed technically, and will be seen to be so through review. Concepts that survive initial review must be refined and re-evaluated through independent studies, leading to useful revision or rejection. We hope that ideas

will spur research as they expose weaknesses in our understanding of the Earth system, as seems to be occurring in the case of oceanic iron fertilization⁵⁻⁷.

Social, ethical and legal questions will surround any serious proposal for global environmental engineering. Incomplete understanding almost always characterizes the behaviour of environmental systems; this is especially true of systems in which living organisms are central, as in the photic zone of the oceans⁵⁻⁷. Without understanding the mechanics of a natural system, there is little basis for predicting the system's response to perturbations. Accordingly, risks are inherent in any large-scale intervention — risks of unanticipated side-effects, of wasting time and resources, or of diverting attention from the root causes of environmental problems.

Who must make decisions and with what criteria? Are there any existing international laws that clearly forbid or regulate the treatment of the Antarctic stratosphere with alkanes or the southern oceans with iron? In any case, we believe that international, but non-governmental, bodies of scientists must be involved in the initial evaluations. Political judgements will also be required to express differing values and perceptions of risks — for example, those most threatened by inaction might press for intervention long before others would be receptive to such ideas.

We hope to stimulate debate and scientific research on these issues; their resolution will require serious, innovative thinking and extensive discussions. Their complexity could demand decades for significant progress. In the meantime we must address the underlying causes of environmental problems, if only to buy more time for us to understand the Earth as a system and the consequences of human interactions with it. □

Ralph J. Cicerone and Scott Elliott are in the Department of Geosciences, University of California, Irvine, California 92717; Richard P. Turco is in the Department of Atmospheric Sciences, University of California, Los Angeles, California 90024, USA.

1. Smith, R. C. *et al.* *Science* **255**, 952–959 (1992).
2. US Patent No. 5,003,186 *Stratospheric Welsbach Seeding for Reduction of Global Warming* (1991).
3. Cicerone, R. J. *et al.* *Science* **254**, 1191–1194 (1991).
4. Prather, M. J. *Nature* **355**, 534–537 (1992).
5. Martin, J. H., Fitzwater, S. E. & Gordon, R. M. *Global biogeochem. Cycles* **4**, 5 (1990).
6. Peng, T. H. & Broecker, W. S. *Nature* **349**, 227 (1990).
7. Sarmiento, J. L. *Global biogeochem. Cycles* **5**, 1 (1991).

Drug-resistant TB may bring epidemic

Research opportunities are opening up for the long-neglected study of *mycobacterium tuberculosis*, a disease of poverty that is once again on the rise. TB is back.

MYCOBACTERIUM tuberculosis (*M. tb.*), a scourge that industrialized nations thought had been eradicated, is resurgent, a fresh reminder that mankind has yet to conquer the bacterial world. Even though TB has infected one-third of the people on this planet, and kills three million every year, it had been thought that there was an armamentarium of antibiotics large and powerful enough to kill any form of the tubercle bacillus that came along.

But within the past three or four years, public health officials have reported an increased incidence of TB mutants resistant to one or more antibiotics. The record is a mutant that is resistant to eleven drugs. In New York City, where the emerging crisis is at its worst, among hospitalized patients in November 1991, 46 per cent of TB isolates were resistant to one or more antibiotics. And therein lies the spectre of an epidemic.

Poverty, homelessness and crowding, which are endemic in New York, Los Angeles and other urban areas, are a new breeding ground for the tubercle bacillus in developed countries, just as has always been the case in the poorer countries of the world where TB is, even today, a major cause of disease and death. The spread of the human immunodeficiency virus (HIV), with its capacity to incapacitate the immune system, has simply added to the host range for TB which is spreading quickly among people with AIDS.

As Barry Bloom, a Howard Hughes Medical Institute investigator at the Albert Einstein College of Medicine in the Bronx observes, "we didn't really know before just how important the link between immune deficiency and TB is". Now we do. And AIDS activists are at the forefront of pressure for new money for treatment and research.

Last week, a committee of the US House of Representatives, chaired by Ted Weiss (Democrat, New York), held hearings to assess the damage caused so far by resurgent TB and to predict the future. It was not an optimistic occasion.

The public health challenge is formidable. Because TB was a major cause of death in Western countries early in this century, a good deal is known about the epidemiology and pathology of the disease. Physicians know, for instance, that TB can be effectively treated in patients who consistently take their medi-

cine for the 6–18 months it can take to destroy all vestiges of *M. tb.* The problem is maintaining therapy in homeless people, drug addicts and others caught in poverty.

The crisis of drug-resistant TB arises because of increasing numbers of people who begin treatment but lapse after a few weeks, allowing the bacillus to evolve protective mutations against antibiotics. Public health authorities told the Weiss committee that the epidemic can be stopped only by a small army of nurses and aides who could go out into the community day after day to stand by while patients take their TB pills. In New York City, where the resurgence of TB appears to be most serious, there are only seven such people to monitor more than 25,000 patients.

The challenge of TB to basic science is no less daunting. In the late 1960s, when TB appeared to be under control, public health programmes were dismantled and TB faded from view. Understandably, it never reached the top of the research agenda in molecular biology during the 1970s and 1980s when diseases such as cancer and then AIDS received the bulk of research funding and attention.

So it is regrettable but no surprise that, from a molecular point of view, the story of *M. tb.* is one of unanswered questions. Furthermore, the pool of researchers in the field is small, as is federal funding. Even in the light of the lurking epidemic, the TB budget at the US National Institutes of Health (NIH) is only \$5.2 million.

The tubercle bacillus has an unusual waxy coat made of up what one investigator describes as "a whole mess of lipids" whose structure and function have yet to be elucidated. Unlike many organisms that multiply rapidly (*Escherichia coli*, for instance), the tubercle bacillus replicates only once every 24 hours. Whereas an *E. coli* colony will form on a plate in about eight hours, *M. tb.* takes about a month to form a colony. Why? It is possible that its slow growth is related to the bacillus's two ribosomal RNA genes, but no one knows that for certain and, in any case, no mechanism of gene regulation is understood.

People have guessed at the mechanisms by which antibiotics effectively kill nonresistant TB strains. Isoniazid, a mainline anti-TB drug, is thought to block

synthesis of mycolic acid in the cell wall. Interference with RNA polymerase is offered as the mechanism by which Rifampin kills the bacillus, but this explanation has not been fully elaborated at a molecular level.

The need for fundamental research on the tubercle bacillus is impeded not only by limited money but also by genuine risk. Unlike the AIDS virus which, lethal though it is, is difficult to contract, people working with TB are all too aware that the bacillus is transmitted by air. At least four physicians in New York are dying of multiple drug-resistant TB. The upshot is that P-3 containment facilities will be absolutely essential. Bloom has a new one, thanks to the Hughes Medical Institute, which built it at a cost of approximately \$180,000. And the US Centers for Disease Control (CDC) are equipped to handle the bacillus, as are a few other laboratories, more are needed.

Testifying before the House committee, CDC director William L. Roper reported progress in the rapid diagnosis of TB, which currently takes 2–8 weeks, but new tests are still only prototypes. DNA amplification techniques that can detect the bacillus in sputum within 48 hours are being tested, and diagnosis with high-performance liquid chromatography within just four hours has been demonstrated in CDC's research laboratories. Studies to identify linkages between TB strains using restriction fragment length polymorphism (RFLP) techniques are also taking place (more than 500 strains from 29 TB clusters have been identified during the past nine months). And CDC, in collaboration with pharmaceutical companies, are screening for new anti-TB drugs.

The real opportunity in this public health crisis is in basic science, where there is plenty of room for young scientists with training in molecular genetics and related fields and where, perhaps, there may be more money. The Human Genome Project, which initially excluded pathogenic organisms, is now funding analysis of *M. tb.*, and its mycobacterial relative, *M. leprae*. The goal is to sequence the entire genome within three to five years, enabling researchers then to identify all the proteins, enzymes, antigens and drug targets in these lethal organisms. But it is likely to be a long haul.

Barbara J. Culliton

A case of missing metals

Howard E. Bond

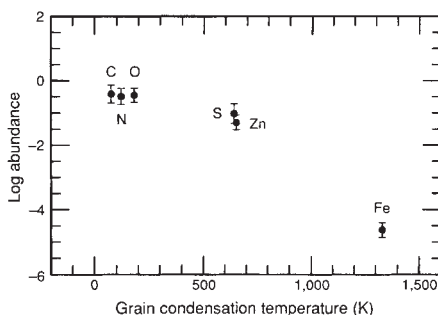
THE chemical composition of a stellar surface is usually assumed to be representative of the material from which the star formed. This assumption gives us the hope of reconstructing the chemical history of our Galaxy by studying the abundances of the elements in stars of various ages. In particular, it is found that the oldest stars are very deficient in 'metals' (chemical elements heavier than helium), and this is interpreted as indicating that the oldest stars were formed before nucleosynthesis in supernovae had a chance to pollute the interstellar medium with heavy elements. But Van Winckel *et al.*, on page 500 of this issue¹, provide the strongest evidence yet that, in at least one class of apparently extremely metal-deficient stars, the deficiencies are due instead to a process that has selectively removed heavy elements from the stellar photospheres — the observed compositions of these objects have nothing to do with any effect of galactic chemical evolution.

Several recognizably peculiar classes of stars are known to have anomalous chemical abundances, owing either to element separation through gravitational settling that can operate in stellar atmospheres under certain restrictive circumstances, or to nuclear reactions inside the star with the processed material being mixed subsequently up to the stellar surface. For the vast majority of ordinary main-sequence and red-giant stars, however, it has been regarded as a safe assumption that little significant modification of the metal abundances has occurred since the star was formed.

Van Winckel *et al.*¹ deal with the star HD52961, a member of the class of post-asymptotic-giant-branch (post-AGB) stars. These are objects that have spent billions of years as main-sequence stars, have recently completed their red-giant evolution and have now commenced a stage of contraction and increasing surface temperature that will, over an interval of 10,000 years or so, transform them into white dwarfs. Members of the class can be recognized as luminous stars lying at considerable distances from the plane of our Galaxy. In the AGB stage, stars throw out dense stellar winds. Post-AGB stars are typically surrounded by circumstellar dust, which is revealed by radiation in the infrared region of the spectrum, and which must have formed by condensation of small solid particles from material in the AGB wind as it cooled while receding from the star.

Several post-AGB stars with moderate to extreme metal deficiencies in their

photospheres have been found in recent years. Examples are HD46703, BD+39° 4926, HD52961 and HR4049, which are deficient in iron (relative to its abundance in the Sun) by factors of about 40, 125, 40,000 and 65,000, respectively^{2–6}. For the first two stars, the metal deficiencies are typical of those seen in the oldest stars. HD52961 and HR4049 have deficiencies among the lowest ever seen in any star in the Milky Way, which could indicate that these two objects are members of Population III, the long-



Abundances (by numbers of atoms, relative to solar abundances; from refs 1 and 6) of six species in the photosphere of the post-AGB star HD52961, plotted against the grain condensation temperature for each species. The correlation indicates that the very low abundance of iron, and the large ratios of C, N, O and S to Fe, are due to a process of depletion onto solid particles that occurred above the photosphere at temperatures between about 700 and 1,300 K, followed by re-accretion of the depleted gas onto the photosphere. The high abundance of zinc is the crucial observation: had the low iron abundance been a consequence of formation of the star early in galactic history, with C, N, O and S the products of nuclear reactions in the interior of the star, zinc would have been equally deficient.

sought primordial stars formed at the beginning of galactic chemical evolution.

An even more bizarre phenomenon, first suspected⁴ in BD+39° 4926, clearly seen³ in HD46703, and subsequently found^{5–7} in HD52961 and HR4049, is an overabundance of sulphur; the S/Fe ratios in these stars exceed those in the Sun by factors of 20–4,000. One could speculate³ that the sulphur was synthesized in the stellar interior and then mixed to the surface. The fact that carbon, nitrogen and oxygen — known to be synthesized in stellar interiors — are also overabundant in these stars might appear to support this possibility but, according to standard theory, the interiors of these low-mass stars never achieve the temperatures required for sulphur synthesis.

Why should extreme metal deficiency be common among post-AGB stars, but

extremely rare among, for example, ordinary red giants⁸; and what does the excess sulphur tell us? Following an earlier suggestion⁹ in a different context, it was proposed¹⁰ that in fact the C, N, O and S abundances in post-AGB stars represent the original stellar composition, and that the iron-group metals have been mostly removed from the stellar photospheres. Indeed, a very similar abundance pattern is seen in the gaseous phase of the interstellar medium¹¹, where it is attributed to depletion of metallic species from the gaseous phase onto solid particles. Relatively volatile species like C, N, O and S form grains only at considerably lower temperatures than do the heavier metals, so that they remain largely as free atoms in the interstellar gas.

A critical test of this explanation is possible, because zinc, unlike the other metals in the iron group, has a low grain condensation temperature. If C, N, O and S are overabundant in post-AGB stars because of nucleosynthesis of light elements, then zinc will share the extreme deficiencies of the iron group. But if the iron group is deficient because of depletion onto grains, then zinc will share the high abundances of the light elements. It is precisely this critical observation that has now been made by Van Winckel *et al.*¹

The result is presented in the accompanying figure, in which the chemical abundances^{1,6} for six species in HD52961 are plotted against the temperatures at which they develop into grains under conditions typical of stellar winds (ref. 12 and references therein). Most of the metallic elements of the iron group cannot be plotted, because they are so deficient that no detectable features are present in the photospheric spectrum of HD52961. However, the spectra obtained by Van Winckel *et al.* show that zinc is as abundant as sulphur. This is dramatic confirmation of the element-depletion explanation.

The problem with this result is that the

1. Van Winckel, H., Mathis, J. S. & Waelkens, C. *Nature* **356**, 500–501 (1992).
2. Luck, R. E. & Bond, H. E. *Astrophys. J.* **279**, 729–737 (1984).
3. Bond, H. E. & Luck, R. E. *Astrophys. J.* **312**, 203–212 (1987).
4. Kodaira, K., Greenstein, J. L. & Oke, J. B. *Astrophys. J.* **159**, 485–512 (1970).
5. Lambert, D. L., Hinkle, K. H. & Luck, R. E. *Astrophys. J.* **333**, 917–924 (1988).
6. Waelkens, C., Van Winckel, H., Bogaert, E. & Trams, N. *R. Astr. Astrophys.* **251**, 495–504 (1991).
7. Takada-Hidai, M. *Pub. Astr. Soc. Pacific* **102**, 139–143 (1990).
8. Bond, H. E. *Astrophys. J.* **248**, 606–611 (1981).
9. Venn, K. A. & Lambert, D. L. *Astrophys. J.* **363**, 234–244 (1990).
10. Bond, H. E. in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 341–349 (Kluwer, Dordrecht, 1991).
11. Spitzer, L. *Physical Processes in the Interstellar Medium*, 206 (Wiley, New York, 1978).
12. Jenkins, E. B. in *Interstellar Processes* (eds Hollenbach, D. J. & Thronson, H. A.) 553–560 (Reidel, Dordrecht, 1987).

temperature in the photosphere of HD52961 is about 6,000–6,500 K, far too high for solid particles to develop. Van Winckel *et al.* suggest a picture in which material left the star during the AGB wind phase, cooled to temperatures below the iron-condensation temperature, and formed metallic grains. Since then, some of the depleted gas (but not the solid particles) has fallen back onto the star, producing a coating of extremely iron-deficient material. One prediction (which fits existing facts, but is open to further tests) is that this phenomenon should be seen only in post-AGB stars with surface temperatures above about 6,000 K; below that temperature, the

stars have convective outer layers that would quickly dilute the accreted depleted material with stellar material of normal composition.

Those who would use metal-deficient stars to construct a history of our Galaxy can probably breathe a sigh of relief. As long as we confine our attention to pre-AGB stars that have not undergone extensive mass loss, dust formation and re-accretion, there is still every reason to think that we can observe the unaltered original compositions. □

Howard E. Bond is at the Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA.

EVOLUTIONARY BIOLOGY

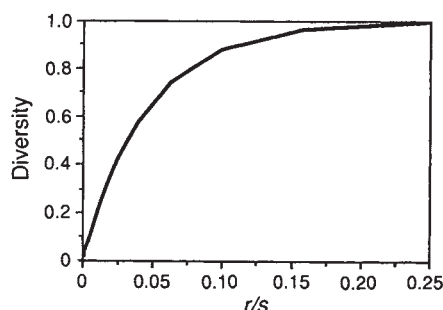
New genes sweep clean

Brian Charlesworth

ON page 519 of this issue¹, Begun and Aquadro present their compilation of data on DNA variation at a large number of loci in *Drosophila melanogaster*, and (where available) on the degree of divergence per nucleotide site at these loci from the sibling species *D. simulans*. They show that there is a statistically significant positive correlation between the intra-specific nucleotide site diversity for a locus and a measure of the frequency of crossing-over per nucleotide in the region in which the gene is located (the coefficient of exchange). No such relation can be detected between inter-species divergence and the coefficient of exchange. Begun and Aquadro interpret these findings in terms of the 'hitch-hiking' process, the consequences of which for molecular variation were first analysed by Maynard Smith and Haigh².

Hitch-hiking is the effect of the spread of a selectively favoured mutation through a population on variation at closely linked sites (it is convenient to assume that this variation is selectively neutral). Consider first a case when there is no recombination. If the successful mutation originates in a single chromosome, its fixation in the population will inevitably cause the simultaneous fixation of all the variants present in that chromosome. Thus, the linked sites effectively go through a bottleneck of population size of one, with the elimination of all variation. The effect of such a 'selective sweep'³ will be detectable if the time elapsed since the sweep is substantially less than $1/u$, where u is the rate of neutral mutation per nucleotide. If there is some recombination between the neutral sites and the selected locus, there is an opportunity for the introduction of genetic diversity at the neutral sites into the gametes carrying the selectively favoured allele

while it is sweeping through the population, so that variability is not so severely reduced (see figure). The expected effect of hitch-hiking on variation in the presence of recombination is one of the more intricate problems in theoretical population genetics. It depends in a complex way on the strength of selection, the frequency of recombination, the population size, and the frequency at which selective sweeps occur in the genome^{2,4–6}. A negative relation between the frequency of crossing-over and the effects of selective sweeps on variability is, however, expected under a wide variety of circumstances^{2,4–6} (see figure).



The effect of recombination on the impact of a selective sweep on genetic variation. The y axis is the expected level of gene diversity at a neutral locus immediately after a single hitch-hiking event in a random-mating population, relative to its initial value. The x axis is the ratio r/s , where r is the frequency of recombination between the neutral locus and the selected locus and s is the fitness advantage to the homozygote for the favoured allele at the selected locus, measured relative to the fitness of the homozygote for the initial allele. The heterozygote is assumed to have intermediate fitness. A large population size (0.5×10^6) and moderate selection ($s = 10^{-3}$) are also assumed. The numerical values were kindly supplied by Wolfgang Stephan.

Observations of unusually low levels of nucleotide site variation for *Drosophila* genes known to be in regions of restricted recombination, such as the tip of the X chromosome, the bases of the major chromosomes and the small, non-recombining fourth chromosome^{1,3,7,8}, have often been interpreted as evidence for selective sweeps. Begun and Aquadro's survey is consistent with the prediction that reduced rates of crossing-over should be associated with reduced genetic variation within a species; the fact that recombination appears to have no effect on inter-species divergence seems to eliminate artefactual explanations, such as a mutagenic effect of crossing-over.

Of course, the data are not without their limitations. Many of the results are obtained from restriction mapping rather than sequencing, so that frequently no distinction can be made between silent nucleotide substitutions that have no effect on amino-acid sequence and ones that result in amino-acid replacements (the former being more likely to fit the neutral assumption⁹). In addition, there is a good deal of variation in the level of both intra-specific and inter-specific variation for a given coefficient of exchange. This makes it difficult to discern details of pattern, for example whether the effect of recombination on variability is most marked for low rates of crossing-over, as predicted by the theory^{5,6} (see figure).

A crucial question concerns the frequency with which selective sweeps have to be postulated in order to explain the observations. This is hard to answer in general, but regions of zero recombination provide some useful insights, although the parameters of interest cannot be estimated very accurately³. Sequence analysis of the fourth chromosome locus *ci^D* showed that ten genes sampled from *D. melanogaster* failed to show any variants among 331 silent nucleotide sites; one variant was detected among nine genes sampled from *D. simulans*³. If this lack of variation is indeed due to a selective sweep, a lower bound to the number of generations t that have elapsed since the sweep is set by assuming that the genes in the sample have descended by completely independent paths from their common ancestor since the sweep³. The probability of observing S mutations among a total of k nucleotides sampled is then given by the Poisson distribution with mean kut . In the case of *D. melanogaster*, the probability of observing no mutations in the sample is $\exp(-3,310ut)$. Using the standard value of $u = 10^{-9}$, a probability of 0.5 of obtaining this result is given by $t = 2.1 \times 10^5$; a similar calculation for the *D. simulans* data yields $t = 5.6 \times 10^5$. This implies that a relatively long period of

time may have elapsed since the postulated sweeps in each species³.

Chromosome four constitutes about 1 per cent of the *Drosophila* genome, so these estimates of t mean that one sweep may occur in the genome every 2,000 generations, or even less frequently. This is much less than Haldane's¹⁰ upper limit to the rate of adaptive evolution (one gene substitution every 300 generations) necessary to avoid extinction due to too high an incidence of genetic death. It is uncertain whether selection at the level of the individual or meiotic drive is responsible for the sweeps, but the evidence compiled by Begun and Aquadro certainly points to the operation of some form of selection leading to allelic fixations in the *Drosophila* genome.

Although more data are desirable to put these conclusions on a firmer basis, they suggest strongly that reconstructions of evolutionary history from gene genealogies may lead to erroneous conclusions about the timing of events at the CLIMATE

population level if the possibility of selective sweeps is ignored, particularly for genomic regions (such as the mitochondrion) where there is no recombination. This is yet another reason for conservatism in the use of such information in interpreting human evolutionary history¹¹. □

Brian Charlesworth is in the Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.

1. Begun, D. J. & Aquadro, C. F. *Nature* **356**, 519–520 (1992).
2. Maynard Smith, J. & Haigh, J. *Genet. Res.* **23**, 23–35 (1974).
3. Berry, A. J., Ajioka, J. W. & Kreitman, M. *Genetics* **129**, 1111–1117 (1991).
4. Thomson, G. *Genetics* **85**, 753–788 (1977).
5. Kaplan, N. L., Hudson, R. R. & Langley, C. H. *Genetics* **123**, 887–899 (1989).
6. Stephan, W., Wiehe, T. H. E. & Lenz, M. W. *Theor. pop. Biol.* **41**, 237–254 (1992).
7. Stephan, W. & Langley, C. H. *Genetics* **121**, 89–99 (1989).
8. Aguade, M., Miyashita, N. & Langley, C. H. *Genetics* **122**, 607–615 (1989).
9. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge University Press, 1983).
10. Haldane, J. B. S. *J. Genet.* **55**, 511–524 (1957).
11. Gee, H. *Nature* **355**, 583 (1992).

Predicting El Niño events

C. F. Ropelewski

It may be possible to predict more accurately the onset of occasional El Niño climate switches, which affect the weather around the whole Pacific basin and beyond. The latest in a series of statistical studies is one by Angell¹ who has noted a correlation with quasibiennial oscillations in the pattern of circulation in the stratosphere. Indeed, Angell (who prepared his paper early last summer) predicted the current deep El Niño event which is bringing extreme drought to northern Australia, Indonesia and southern Africa and wet conditions to parts of North and South America.

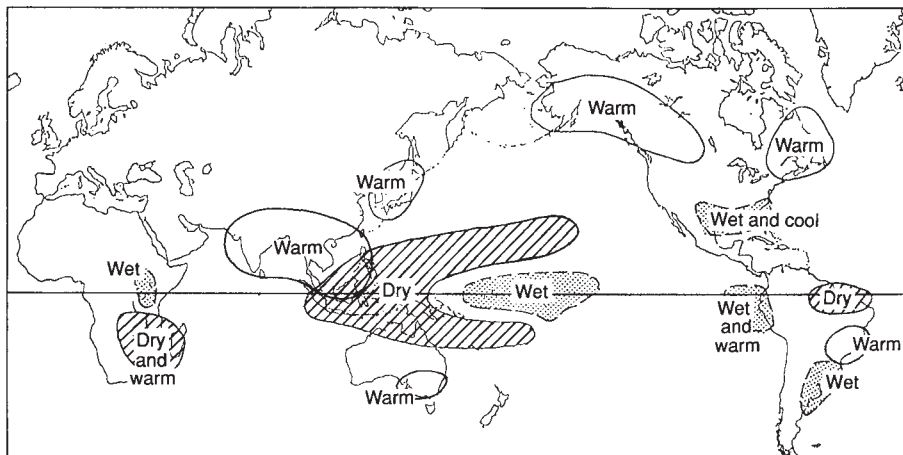
Angell's study fits into a trend of increased understanding of short-term climate variability that has been progressing for two decades. The El Niño ('the boy child' so-named by Peruvians from its onset around Christmas) and the related Southern Oscillation, known collectively as ENSO, is a quasiperiodic variation in climate which occurs every two to seven years. It is associated with significant changes in normal rain and temperature patterns over large part of the globe (see figure). Because individual ENSO episodes can last one to two years, the mere identification of a current event provides forecasters with a powerful tool for predicting seasonal climate for certain regions. But because ENSO's have predictable seasonal patterns of precipitation and temperature, to foresee an ENSO event would be even more useful².

The ENSO phenomenon results from a complex interaction between the tropical Pacific ocean and the atmosphere. It involves a quasiperiodic waxing and waning of atmospheric circulation, best indexed by the strength of the equatorial easterly trade winds, and the ebb and flow of oceanic density structures and currents, best indexed by the appearance and disappearance of above-normal sea surface temperatures in the equatorial central and eastern Pacific. The timescales associated with the lifetime of a particular ENSO (one to two years) and the timescales associated with the periods between ENSO episodes (generally two to seven years) are thought to

be constrained by the physical boundaries of the Pacific Ocean basin and by the timescales associated with interactions between ocean and atmosphere in this coupled, two-fluid system.

Most ENSO research has concentrated on understanding ways in which energy is transferred at the ocean-atmosphere interface and on understanding how these interactions relate to changes in the global atmospheric circulation. Links between the ENSO and the stratosphere have been suggested previously³, but this new study by Angell¹ and independent related work by Gray and colleagues⁴ are the first to suggest that the stratospheric quasibiennial oscillation (QBO) may hold a key to ENSO prediction. In particular, Angell proposes that the relationship is not one-to-one, and that it may "skip a beat" after a particularly strong ENSO, so explaining the irregularity of the phenomenon.

The stratospheric QBO, discovered in the early 1960s, is most clearly seen as a fairly regular oscillation (mean period near 26 months) in the winds at heights between 20 and 30 km. Although the workings of the stratospheric QBO are quite complex, and some details are still debated, the most widely accepted explanations centre on the upward propagation of energy associated with internal 'gravity' waves (which are similar to more familiar ocean waves but occur in the atmosphere across density surfaces). It has previously been suggested that the stratospheric QBO could influence the lower atmosphere, but these ideas have generally been dismissed on the grounds that the stratosphere is much too small, a mere tail, to wag the dog of global tropospheric circulation. Angell presents no physical mechanism to explain the link between the stratospheric QBO and the ENSO, but suggests that it may be related to the strength of the upper branch of the Walker circulation. This circulation is a direct east-west thermal cell whose mean circulation takes place



Typical rainfall and temperature patterns associated with El Niño/Southern Oscillation (ENSO) conditions for the Northern Hemisphere winter season⁹.

in a vertical plane along the equatorial Pacific. Its primary features in the central Pacific are near-surface easterlies and upper-tropospheric westerlies.

To complicate the picture, there is considerable evidence, in tropospheric indices, to suggest a strong component of biennial or near-biennial variability in the timescales associated with the recurrence of ENSO^{5,6}. The links, if any, between tropospheric biennial variability and the stratospheric QBO have not been found to be statistically significant^{2,5}. With stratospheric data available only back to 1954, and no *a priori* physical hypothesis to link ENSO and the stratospheric QBO, there are no tools available to test Angell's hypothesis independently.

The possible links between the ENSO and the stratospheric QBO will, no doubt, be monitored for years to come, but these are not the only ways of predicting ENSO occurrence. There is a body of statistical tools, such as canonical correlation analysis², that use less esoteric predictors such as surface pressure and sea surface temperature, and that have been routinely used⁷ for ENSO prediction with some success.

However, the real future of seasonal climate prediction is more likely to be in the rapidly developing arena of coupled ocean-atmosphere numerical models. To a large extent, seasonal climate prediction has reached a point analogous to the state of numerical weather prediction in the early 1950s. Simple versions of these models have been successful in ENSO prediction (Cane and Zebiak's model⁸ predicted the current ENSO a year in advance), and there is reason to believe that useful multi-season predictions based on numerical models may be available in the near future.

Thus, if the numerical climate models match their potential, the statistical prediction schemes are likely to have short lifetimes as aids to forecasts. Statistical models may find their primary role to be as benchmarks for the computer models to beat or they may be used in combination with the coupled climate models for seasonal rainfall and temperature prediction. □

C. F. Ropelewski is in the Climatic Analysis Center, World Weather Building, National Oceanic and Atmospheric Administration, Washington DC 20233, USA.

1. Angell, J. *Geophys. Res. Lett.* **19**, 285–288 (1992).
2. Barnett, T. *et al. Science* **241**, 192–196 (1988).
3. Yasunari, T. *J. met. Soc. Japan* **67**, 483–493 (1989).
4. Gray, W. M., Sheaffer, J. D. & Knagg, J. A. *Geophys. Res. Lett.* **19**, 107–110 (1992).
5. Barnett, T. *J. Climate* **4**, 269–285 (1991).
6. Rasmusson, E. M., Wang, X. & Ropelewski, C. F. *J. Mar. Syst.* **1**, 71–96 (1990).
7. Barnston, A. G. & Ropelewski, C. F. *J. Climate* (in the press).
8. Zebiak, S. E. & Cane, M. A. *Mon. Weath. Rev.* **115**, 2262–2278 (1987).
9. Ropelewski, C. F. & Halpert, M. S. *Mon. Weath. Rev.* **115**, 1606–1626 (1987).

Making and breaking bonds

A. W. Kleyn

THE crucial role played by excitation in the disintegration of molecules at solid surfaces — often the first step in heterogeneous catalysis — is revealed in experiments reported by Hodgson *et al.* on page 501 of this issue¹. The authors followed the fate of hydrogen molecules fired at a sample of copper, and found that faster-moving molecules, or those whose internal bond was vibrationally excited, are more likely to break up and stick to the surface.

There is, in fact, little understanding of the pathways leading to the breaking and making of chemical bonds at surfaces, which is central to many important processes, such as the fabrication of semiconductor devices, surface reactivity in catalysis, and friction and adhesion on a nanoscopic scale. As the ultimate engineering of surface effects will be carried out at the atomic level, a fundamental understanding of the events occurring at this level is indispensable. It is not that such things have not been studied until now. On the contrary, even in the 1930s Lennard-Jones and Devonshire worked on bond breaking or molecular dissociation at surfaces, realizing that there are two modes of interaction between a molecule and a surface — the weak Van der Waals force of the intact molecule, and the stronger chemical bonding of the constituent atoms following dissociation at the surface (Fig. 1). This picture neglects the internal degrees of freedom of the solid and the effect of molecular orientation, which is not necessarily a shortcoming for most purposes, but hints at the one vital degree of freedom — the internuclear separation in the molecule — which is included in all thorough modern treat-

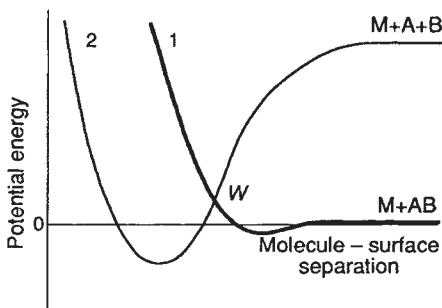


FIG. 1 Potential energy diagram for the approach of a molecule (AB) towards a surface (M), as envisaged in the 1930s by Lennard-Jones. Curve 1 represents the weak Van der Waals interaction of the undissociated molecule with the surface, and curve 2 the stronger chemical interaction of the dissociated molecule (individual atoms). Note that dissociation at the surface releases energy, but that a potential barrier *W* has to be overcome for this to happen.

ments of such systems. The two curves in Fig. 1 represent the two extremes in the two-dimensional potential energy hypersurfaces used nowadays, which are contour plots of the interaction as a function of internuclear and molecule-surface separation (for example, Fig. 4 of Hodgson *et al.* on page 503): the first curve

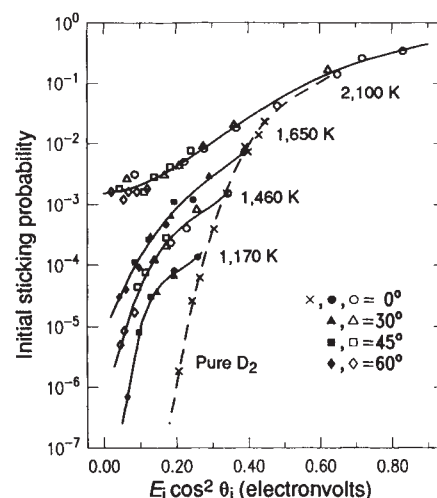


FIG. 2 Sticking probability for D_2 molecules fired at an angle θ at the copper (111) surface. Data for seeded beams at four temperatures (solid lines) and for pure D_2 at 875–2,100 K (broken line) are shown. E_i , collision energy. (From ref. 4.)

is a slice through the surface at equilibrium internuclear separation, the other a representation with widely separated atoms.

What is clear from either depiction is that energy is gained by breaking the molecular bond. This may result in a violent rupture of this bond, leading to rapid separation of the molecule's constituent atoms. A vivid illustration of this rupture was recently reported by Brune *et al.*², who show that the oxygen atoms resulting from dissociation of thermal O_2 impinging at an aluminium surface travel at least 40 Å over the surface. The observations of this process were made with a static probe, long after the dissociation event, with a scanning tunnelling microscope (STM). Information about the dynamics of the dissociation event cannot be taken directly from this STM observation.

The presence of a potential energy barrier, which has to be surmounted by the approaching molecules, is also evident in Lennard-Jones's diagram. It is the nature of this barrier that is investigated by Hodgson *et al.* in this issue. The authors use indirect methods, as molecular dissociation at surfaces cannot yet be probed directly by the time-dependent

laser techniques that have been developed for gas-phase reactions (see I. W. M. Smith's News and Views article³). The measurement of the sticking coefficient as a function of the molecule's

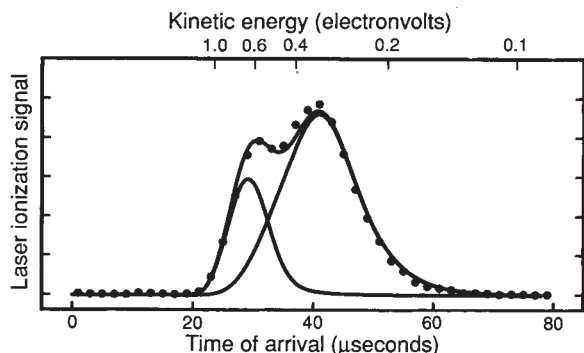


FIG. 3 Time of arrival of H_2 molecules in a pulsed beam scattered from Cu(111) and detected in the $v = 1$ vibrational state. Two components are apparent: to the right, the slower $v = 1$ molecules that failed to stick to the copper surface; to the left, faster molecules initially in the $v = 0$ state, which have converted some kinetic energy into vibrational excitation. The authors also found a clear dependence on rotational excitation. (From ref. 5, courtesy of D. J. Auerbach.)

translational and internal energy, or as a function of the surface conditions, surfaces instead. In such experiments, information on the dynamics of the reaction (how the process occurs on a femto-second timescale) is lost. Nevertheless, vital clues can be retrieved.

The reaction of deuterium molecules (D_2) with the copper (111) surface is the fruitfully of surface chemistry studies. Only recently, Rettner *et al.*⁴ remeasured the sticking probability of D_2 on Cu(111) as a function of the kinetic energy corresponding to the velocity normal to the surface (Fig. 2). They did this by firing a supersonic beam of D_2 molecules, carried in a stream of carrier gas, at the copper surface at various angles and detecting those that stuck. The fact that it is only the energy of motion normal to the surface that matters is borne out by the similarity (apart from scaling) between results with beams at different angles. This indifference to motion parallel to the surface is to be expected for a close-packed and, hence, smooth surface like Cu(111).

Rettner *et al.* controlled the translational and vibrational energy of the D_2 molecules by altering the temperature of the nozzle source of their beam and by altering the carrier gas. With a pure D_2 beam, the kinetic and vibrational energy vary in step with temperature. By diluting the beam in H_2 or rare gases, the authors could alter the vibrational energy (dependent only on the nozzle temperature) separately from the translational energy. In this way, they found that for a given translational energy, an increase in the vibrational energy gives an enormous increase in the sticking

probability. Clearly, vibrational energy helps the molecule to surmount the energy barrier. This in turn indicates that the barrier (as depicted on the potential-energy hypersurface) is located along the vibrational (or bond-stretching) coordinate, and not along the molecule-surface separation coordinate (see Fig. 4 of Hodgson *et al.*).

Hodgson *et al.*, and also Rettner *et al.* in work yet to appear⁵, now take this work one step further, verifying directly the effect of vibrational excitation. They do this by selectively detecting vibrationally excited molecules reflected from the surface; the technique of multiphoton ionization of the scattered molecules allows them to discriminate between those in the vibrational ground state ($v = 0$) from those with a single vibrational quantum ($v = 1$). Hodgson *et al.* find that as soon as the translational energy of the molecular beam is high enough (at about 0.35 electronvolts) for the excited molecules to overcome the barrier, far fewer of these are reflected, owing to vibrationally promoted sticking.

But these experiments can do more, because the molecular vibrations effectively provide a natural internal clock

that runs at the timescale of the collisional dissociation process. Both groups find that molecules scattered from the copper surface (those that fail to stick) can become vibrationally excited during the collision: at higher translational energy, the population of elastically reflected $v = 1$ molecules is supplemented by others whose translational energy has been partially converted into vibration. The two components can be distinguished because their different velocities ensure that they arrive (when the beam is pulsed) at the detector at different times (Fig. 3 in this article and Fig. 3 in Hodgson *et al.*, page 503).

Because both vibrational excitation and molecular dissociation involve stretching a molecular bond, the latter beyond the limits of the bond strength, this close alliance between sticking probabilities and collisional excitation makes sense. But these new reports describe the first experiments to identify molecules that nearly reacted at a surface, but failed to do so. □

A. W. Kleyn is at the FOM Institute for Atomic and Molecular Physics, Kruislaan 407, 1098 SJ Amsterdam, The Netherlands.

1. Hodgson, A., Moryl, J., Traversaro, P. & Zhao, H. *Nature* **356**, 501–504 (1992).
2. Brune, H., Winterlin, J., Behm, R. J. & Ertl, G. *Phys. Rev. Lett.* **68**, 624–626 (1992).
3. Smith, I. W. M. *Nature* **343**, 691–692 (1990).
4. Rettner, C. T., Auerbach, D. J. & Michelsen, H. A. *Phys. Rev. Lett.* **68**, 1164–1167 (1992).
5. Rettner, C. T., Auerbach, D. J. & Michelsen, H. A. *Phys. Rev. Lett.* (in the press).

PROTEIN STRUCTURE

New paths from dead ends

Willie Taylor

A THREE-dimensional model of a protein can be constructed for an amino-acid sequence, provided that the sequence shares some similarity with another sequence of known structure. This activity, commonly called modelling by homology, allows a limited amount of hard-won structural data to be extrapolated with more easily obtained sequence data. As the sequence and structure databanks expand it is becoming apparent that, relative to the number of sequences, nature's repertoire of structures may be quite limited. The construction of protein models is, therefore, likely to become an increasingly important technique, and any computational aids to speed it along are always welcome.

On page 539 of this issue¹, Desmet *et al.* present a new algorithm that promises to both speed and improve an aspect of the modelling process. They tackle the problem of how to pack the amino-acid side-chains of the new sequence.

Equating the overall modelling process with a game of chess, this stage corresponds to the end-game, being both last and in places quite forced, because once a key residue has been moved into position it can greatly restrict the conformational options of its neighbours. But this does not mean that the problem is trivial — as with end-games in chess, there is effectively an infinite number of combinations from which one (the fastest win) must be selected.

The preceding and less-well-characterized aspects of the modelling game involve finding the initial sequence similarity (which may be fragmentary) and establishing an alignment over the length of the sequences. The best alignment will often require relative insertions and deletions of sequence, which in structural terms imply that the main-chain (backbone) of the given structure must be modified. Avoiding these difficulties, Desmet *et al.* begin with a fixed backbone structure on which a new set

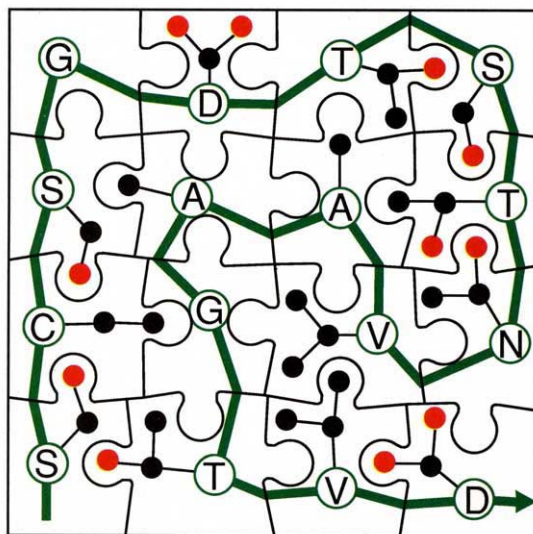
of side-chains is to be added. For relatively similar sequences this is not an unreasonable starting point, as it is clear from many examples that protein structures are better conserved through evolution than the sequences which overlay them.

Early protein models were built 'by hand' using interactive computer graphics^{2,3}, and their quality depended largely on the expertise of the modeller. With reasonably related sequences, however, artistic freedom in the positioning of side-chains can be cramped. Clearly, as a first approximation, identical residues in the two proteins should retain their known conformation; and if the new amino acid is fully or partially isosteric with its equivalent in the known structure, then the positions of corresponding atoms should be retained as far as possible. Conflicts and clashes inevitably arise, but they can usually be resolved by energy minimization, often with very little main-chain movement, resulting in quite plausible protein structures.

Some attempts to automate this process have relied on the observation that side-chains do not adopt all imaginable positions, and that as few as 74 different conformations (referred to as rotamers) might be sufficient⁴. Using these discrete conformations, the computational problem becomes one of choosing a set that fits together. This is rather like a jigsaw puzzle in which the protein structure is the picture and the amino acids (or residues) are the pieces. The added twist is that each piece has a number of different forms and the problem is to find the combination of them that will create a fully interlocked picture (see figure). Faced with this puzzle, a natural approach would be to start at any point and extend out by trial-and-error, undoing any dead ends until another avenue of addition opens up. In computer terms this would be called a combinatorial tree search and, depending on how good one became at spotting dead ends, it could be said that a branch-and-bound algorithm was being used. Yet despite such sophisticated terminology, it would still take an enormously long time to solve a puzzle of a size comparable to packing the side-chains of a small protein.

This seemingly intractable combinatorial problem has been reduced by Desmet *et al.* to a manageable size by the use of a simple logical device, which they call dead-end elimination. Their argument is given in terms of energy inequalities, but for simplicity I will (not unreasonably) assume that side-chains

have only two energy states — those that bump into something (either the main-chain, another side-chain or themselves) and those that do not. The second state is, of course, the desired (low-energy) state. Consider now a residue in a protein that has two possible rotamers (R and T). Desmet *et al.* prove that if rotamer R always bumps into something, no matter how you try to move the other side-chains out of its way and, in addition, if rotamer T is always clear, no matter how you try and make the others bump into it, then rotamer R can be



The protein jigsaw puzzle. At first sight the solution is easy because there is a known backbone structure (green) to copy. But packing the side-chains (small red and black circles) is difficult, because for each piece there are a number of alternatives (rotamers) only one of which will appear in the completed picture at any position. The approach of Desmet *et al.* can be explained, in simplified terms, by considering the options for the residue (C) at the second position. If there are three rotamers for C and two rotamers for S, then each C is tried with each S at the first and third positions. If there is a rotamer of C that will not fit with any S at either adjacent position (or with G at the thirteenth position), then that piece cannot be part of the final picture and can be thrown away. This test is applied to all positions, so reducing the number of pieces that need to be considered when it comes to the final (combinatorial) assembly stage.

eliminated as a possible state for that residue. More explicitly, the condition requires that the sum of best-pair interaction energies for R must be greater than the sum of worst energies for T (see equation 2 on page 540). The power of this condition is that its calculation involves only pairs of residues and is therefore of the order of N^2 rather than combinatoric.

Credit must go to Desmet *et al.*, not only for showing the truth of this condition but also for realizing that it would provide a useful constraint on their combinatoric packing problem, because it is not intuitively obvious (to me anyway) why the method should work. When

building side-chains 'by hand' it is all too easy to bend one into the middle of another and, with a twelfth-power repulsive energy term, any large number might be generated arbitrarily (or by chance when using discrete rotamers). Yet it is precisely these values that the truth of the dead-end condition relies upon. It is rather unsettling to think that rotamer elimination might rely on the difference of 'astronomic' numbers.

Nonetheless, the method clearly works. For a small protein of 76 residues, Desmet *et al.*'s results show that repeated application of dead-end elimination to pairs of residues can reduce a problem of 10^{76} possible combinations down to 7,000 (with all but seven side-chains completely determined). Further application of the method to triples and higher combinations then produced a unique solution. The quality of the result (compared to the known X-ray structure) is good, especially for residues in the constrained core. Previous 'conventional' methods can also generate accurate predictions in the core region, and do so in a comparable time^{5,6}. But the advantage of dead-end elimination is that the solution (if unique) must be a global minimum, a result that cannot be guaranteed by stochastic techniques such as the Monte-Carlo method or simulated annealing.

From a theoretical point of view, the new approach is very attractive. In practice, however, its value may be limited by the underlying assumption that the main-chain cannot move. This condition will be only approximately attained when the two sequences are very similar — so similar that most side-chains in the core will be identical or largely isosteric — and in these circumstances it is best to incorporate as much of the known data as one can by simply leaving positions unchanged as far as possible. For more remotely related sequences, the method may find a use in deciding

whether it is necessary to move the main-chain, and if so, to pinpoint the most parsimonious moves to allow a packing solution.

Moving up the structural hierarchy, it

- Desmet, J., De Maeyer, M., Hazes, B. & Lasters, I. *Nature* **356**, 539–542 (1992).
- Greer, J. *J. molec. Biol.* **153**, 1027–1042 (1981).
- Blundell, T. L., Sibanda, B. L. & Pearl, L. H. *Nature* **304**, 273–275 (1983).
- Ponder, J. W. & Richards, F. M. *J. molec. Biol.* **193**, 775–791 (1987).
- Lee, C. & Subbiah, S. *J. molec. Biol.* **217**, 373–388 (1991).
- Holm, L. & Sander, C. *J. molec. Biol.* **218**, 183–194 (1991).
- Chothia, C. *et al. Nature* **342**, 877–883 (1989).
- Wilmot, C. & Thornton, J. M. *J. molec. Biol.* **203**, 221–232 (1988).
- Taylor, W. R. *Protein Eng.* **4**, 853–870 (1991).

is interesting to speculate on whether the method might help with some of the combinatorial problems encountered in protein structure prediction. In this respect, the packing of antibody loops round an antigen-binding site⁷, or simply the packing of loops in general⁸, would seem to be a logical progression because both occur in a limited variety of conformations. The combinatorics of fold prediction by secondary structure packing⁹ might also benefit, but here the constraints are soft and without hard

PARASITOLOGY

Skirmishes on the border

John E. Donelson and Alice B. Fulton

INTRACELLULAR parasites routinely breach the outer membranes of their host cells, gram-negative bacteria of the genus *Yersinia* using a surface protein called invasin for the purpose. In a paper in the *Journal of Cell Biology*¹, Young, Falkow and Schoolnik show that, when the *Yersinia* gene for invasin is cloned in a normally noninvasive strain of *Escherichia coli*, the bacteria can enter human epithelial cells and cytoskeletal proteins reorganize around them. Inert latex beads coated with invasin also enter the cells, demonstrating that other bacterial components are not necessary for the attack to succeed.

Three species of *Yersinia* are human pathogens. *Yersinia pestis* is transmitted incidentally to humans from infected fleas of rodents and has caused outbreaks of plague throughout the ages; *Y. enterocolitica* and *Y. pseudotuberculosis* are acquired from contaminated food or water and cause a variety of gastrointestinal symptoms. Although *Yersinia* can enter mammalian cells by several different mechanisms, the best characterized pathway involves binding of invasin to host-cell integrins, a group of large $\alpha\beta$ heterodimeric proteins that span the membrane and participate in cell-cell adhesion². An earlier study showed that only the C-terminal 20 per cent of invasin is necessary for adherence of *Yersinia* to the target cell and for its passage into it³.

Young and colleagues now find that *E. coli* expressing invasin from *Y. enterocolitica* first attach rapidly and reversibly to single sites on the filopodia of cultured epithelial cells. Within a few minutes the bacteria irreversibly make several points of contact with the epithelial cell surface and are internalized within endosomes. Adherence is blocked by antibodies that bind to the β_1 component

bumps it is unlikely that the dead-end condition would retain its power. Outside the world of protein structure, there are many intractable combinatorial problems. If the algorithm is applicable to even a few of them, it will be of great importance — and may even result in better computer chess. □

Willie Taylor is in the Laboratory of Mathematical Biology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

of integrins, suggesting that invasin interacts with this constituent of the membrane. Immunofluorescence microscopy shows that after the bacteria enter the cell, they are surrounded by polymerized actin as well as actin-associated proteins such as filamin, talin and the β_1 -integrin subunit. From the dramatic accumulation of these cell-cortex proteins around the invasin-expressing *E. coli*, it seems

small radius of curvature, the cell may simply engulf it instead of proceeding over it.

Host cells can fight off parasites by altering their surface receptors. For example, the protozoan malaria parasite uses a transmembrane protein of red blood cells, called band 3, as a receptor for entering a cell. The anion-transporting band 3 protein associates with ankyrin, a protein that links the spectrin-based submembrane skeleton to the red-cell membrane. Liu *et al.*⁴ have shown that the membranes of red blood cells in southeast Asian patients with ovalocytosis possess a structurally and functionally abnormal band 3 protein that has enhanced binding to ankyrin and reduced lateral mobility, resulting in increased rigidity of the red-cell membrane. It has been shown experimentally that these red blood cells resist invasion by malaria parasites, but the biochemical basis of the band 3 modification conferring this resistance has yet to be determined unambiguously^{5,6}. Thus, the host cell can counter a parasite's attack by reinforcing its border, but at a potential cost, in this case reduction of its structural flexibility.

Still other factors can be involved in breaching host-cell membranes. Many parasites subvert their hosts' serum proteins to facilitate their attachment and entry into cells⁷. In the case of *Leishmania* — protozoa that parasitize macrophages — the exploited serum protein is C3, a major constituent of the complement system. Activation of complement is normally a first line of defence against pathogens; indeed, during the initial stage of *Leishmania* attack C3 is covalently bound to molecules on the parasite surface and converted to breakdown products, C3b and iC3b (ref. 8). At least one of these proteins, iC3b, specifically binds to an integrin on the surface of macrophages, variously called Mac-1, CR3 or CD11b/CD18.

When this binding takes place, the parasite can enter a macrophage and multiply within the acidic and hydrolase-rich environment of its phagolysosomes. So *Leishmania* exploits two components of its host's defence mechanisms — complement and macrophages — for its own survival.

Mosser, Springer and Diamond⁹ have examined this process by cloning the complementary DNA for human macrophage Mac-1 into COS cells, a monkey fibroblast cell line that is not susceptible to leishmanial infection. *Leishmania* binds to the transfected COS cells expressing Mac-1 in the presence of complement, but not in its absence. This complement-dependent binding can be

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Thin-section electron micrograph from the report by Young *et al.*¹, showing that a normally noninvasive strain of *Escherichia coli* expressing invasion from *Yersinia enterocolitica* on its surface can attach to human epithelial cells and be internalized within individual endosomes. Arrowheads show points of attachment between the bacterial and cell membranes. Scale bar, 1 μ m.

that a high density of binding sites on the bacteria preferentially attract these proteins away from their usual cellular locations. The localized polymerization of actin around the endosomes appears to be similar to that induced during cell-surface capping, phagocytosis or cell adhesion, and is functionally significant; cytochalasin D, an inhibitor of actin polymerization, prevents entry of the bacteria, indicating that the process requires *de novo* microfilament polymerization. The entire sequence of events resembles the movement of a cultured cell whose filopodia seek out new surfaces and pull the leading edge of the cell body forward. If the new surface is an invasin-covered bacterium with a

mimicked using purified Mac-1 attached to a plastic surface. But binding to the transfected COS cells is not followed by entry of the parasite into the cell, suggesting that, unlike the interaction between *Yersinia* invasins and the β_1 subunit of epithelial integrins, other components of the *Leishmania*/macrophage interaction are necessary for the parasite attack to succeed⁷. After *Leishmania* is bound, its entry into the cell may simply be a part of the macrophages' normal phagocytosis mechanism for eradicating complement-coated foreign material from the bloodstream.

These three studies illustrate different approaches for examining the first steps of the interaction between parasites and host cells. In the case of *Yersinia*, a surface component of the parasite was expressed in another organism; in that of *Leishmania*, a host-cell component was expressed in another cell; and in that of malaria, an aberrant constituent of the

host-cell membrane was characterized. But, informative as they are, the studies show that we have barely scratched the surface in understanding the border skirmishes so important to the survival of both intracellular parasites and their hosts. □

John E. Donelson of the Howard Hughes Medical Institute and Alice B. Fulton are in the Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242, USA.

1. Young, V. B., Falkow, S. & Schoolnik, G. K. *J. Cell Biol.* **116**, 197–207 (1992).
2. Falkow, S. *Cell* **65**, 1099–1102 (1991).
3. Leong, J. M., Fournier, R. & Isberg, R. R. *EMBO J.* **9**, 1979–1989 (1990).
4. Liu, S.-C. *et al. New Engl. J. Med.* **323**, 1530–1538 (1990).
5. Jones, G. L., Edmundson, H. M., Wesche, D. & Saul, A. *Biochim. biophys. Acta* **1096**, 33–40 (1991).
6. Husain-Chishti, A. & Ruff, P. *Biochim. biophys. Acta* **1096**, 263–264 (1991).
7. Hall, B. F. & Joiner, K. A. *Immun. Today* **12**, A22–A27 (1991).
8. Russell, D. G. & Talamas-Rohana, P. *Immun. Today* **10**, 328–332 (1989).
9. Mosser, D. M., Springer, T. A. & Diamond, M. S. *J. Cell Biol.* **116**, 511–520 (1992).

Tracing great fluid migrations

Dimitri A. Sverjensky and Grant Garven

DATING measurements presented by Brannon *et al.* on page 509 of this issue¹, correlating the origin of metal sulphide ores in the Mississippi Valley with an episode of tectonics that built the Appalachian Mountains 700 km away, apparently confirm that chemically altering fluids can be driven through vast regions of the continental interior. The authors use dating methods previously used by Nakai *et al.* two years ago², which related similar Mississippi Valley type (MVT) ores in the Appalachians themselves to a separate episode of mountain building, revealing the ubiquitous nature of such processes.

Hot, saline, aqueous (hydrothermal) fluids circulate throughout the Earth's crust, and at distinct times in the geological past such circulation has been associated with the development of extraordinary concentrations of scarce metals (generally metal sulphides) that constitute the ore deposits on which industrial society depends. Establishing the exact age of metallic ore deposits and the processes associated with their formation has long proved tough. But it is now understood that some classes of ore deposits contain records of continental-scale migration of hydrothermal fluids. These deposits and their genesis are worth studying, therefore, not only for their economic significance but also for what they will reveal of the history and evolution of the Earth's crust. For example, studies of MVT deposits, which are major sources of zinc, lead, barium and fluorine, have tied in mas-

sive hydrothermal fluid-migration events across the mid-continent of the United States with the evolution of the Appalachian and Ouachita mountain belts. But the link is not rock-solid, because the ages of the MVT deposits have not been pinpointed unambiguously. Consequently, speculation about the exact timing of the fluid migrations and their relation to specific tectonic events is just that.

Ore deposits are normally dated by concentrations of radiogenic isotopes using either the ore minerals themselves (such as galena, PbS, or sphalerite, ZnS) or minerals thought to have formed at the same time by the chemical interaction of the hydrothermal fluids with the host rocks (for example, hydrothermal muscovite or K-feldspar). But lead-isotope dating does not work for the deposits in the Mississippi Valley region because they contain lead from several highly varied sources with a range of unknown parent U/Pb ratios. And the K/Ar or Ar/Ar methods yield ambiguous results because there are no well-defined hydrothermal alteration assemblages containing potassium silicates in MVT deposits.

An alternative is to look at the extremely low levels of elements such as rubidium, strontium, samarium or neodymium that are associated with the ore minerals, such as sphalerite (ZnS), in MVT deposits. Elements like rubidium and strontium would be expected to enter the crystal structure of sphalerite with great reluctance, because of their

size and charge. Nevertheless, Nakai *et al.* and Brannon *et al.* show that the isotopic compositions of even these very low levels of Rb and Sr, along with those trapped in fluid inclusions, can be used to obtain the ages of sphalerite samples, and hence the age of migration of the ore-forming fluids. These studies give the most definitive dates yet obtained of ore formation for MVT deposits.

Nakai *et al.* studied sphalerite from an MVT deposit within the Appalachian mountain belt itself (the zinc-rich east Tennessee ore district) and found an age of 377 ± 29 million years. This coincides with the Acadian phase of mountain building in the Appalachians. There are many other (generally small) deposits along the length of the Appalachians from Georgia to Newfoundland that may be of similar age. In contrast, Brannon *et al.* use the same technique on sphalerite from one of the classic deposits in the Mississippi Valley region — the Upper Mississippi Valley district — and find that this formed 270 ± 4 million years ago, at the same time as the Alleghenian phase of mountain building in the Appalachians, even though the two regions are separated by over 700 km.

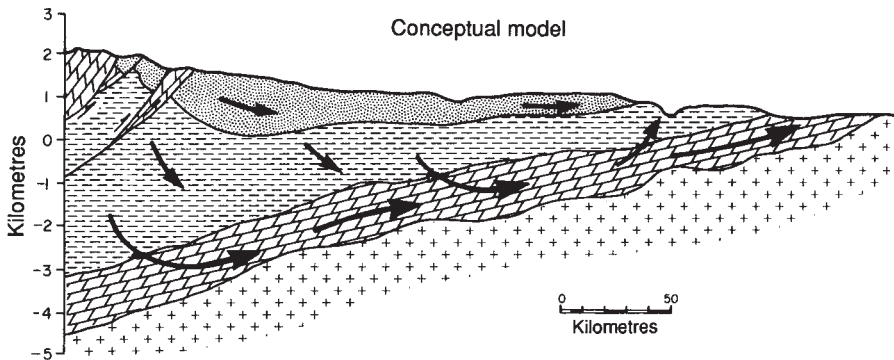
It is already well established that the ore-forming fluids that migrated through deposits in the greater Mississippi Valley region were chemically similar to those found at depths of 3–6 km in sedimentary basins such as the Illinois basin, particularly the brines found in oil-fields³: they typically had temperatures of about 50–220 °C and were rich in sodium and calcium chloride. But how and why did the large-scale migration of hot saline fluid take place? During the past ten years, several models have emerged to address these questions.

One formerly popular model invoked expulsion of basinal brines by compaction. But the flows generated by subsidence compaction seem to be either small in volume or short-lived, and the thermal perturbations created by such flows are unequal to the range of precipitation temperatures recorded in fluid inclusions. Perhaps tectonic compression, thrusting and dilation of the Appalachian and Ouachita mountain belts drove the brine migration to create both oil fields and ore districts^{4,5}. Hydrological calculations^{4,6} of this 'squeeze' effect, however, again suggest that rates of flow associated with compression in an orogenic belt are low — less than 10 cm yr⁻¹. It has also been conjectured that heat may be released catastrophically from the deep crust during orogenesis as a result of free convection of fluids in permeable continental crust⁷. But regional permeabilities in excess of 10⁻¹⁵ m² are required for free convection, conditions that seem unlikely to be

duplicated over continental-scales in the deep basement.

Another possible mechanism involves pumping of fluids along regional faults during seismic events⁸. Seismic pumping can generate flow volumes of the order of $5 \times 10^6 \text{ m}^3$. But it is difficult to see how hot seismic pumping could explain the widespread and pervasive geoche-

vated near the edges of foreland platforms and along intracratonic arches, thereby explaining the typical deposition temperatures reflected by fluid inclusions. In addition, thermal and chemical transients would occur throughout the history of the ore deposition as the hydrological system adjusted to the landscapes generated by mountain building.



Schematic cross-section illustrating the style of gravity-driven, continental-scale fluid migration in uplifted foreland basins, which are formed in front of mountain belts created by compressional tectonics. The arrows indicate the focusing of deep groundwater and brine through a Cambrian-Ordovician carbonate aquifer (slanted brick pattern) which is underlain by nearly impervious Precambrian basement and overlain by low-permeability mudstones. Foreland sedimentary basins such as those in front of the ancient Appalachian and Ouachita Mountains provided the metal-bearing brines that were driven several hundreds of kilometres across the US midcontinent to form large ore deposits in discharge areas along the Mississippi Valley region. The hydraulic force responsible for brine migration is attributable to the newly elevated topography. This gravity-driven flow system decayed with erosion of the landscape, but flow rates of metres per year were capable of elevating heat flow at the basin margin for millions of years. The flow system portrayed here was apparently established after the Alleghenian-Ouachitan phases of mountain building (300–250 million years ago). Earlier periods of mountain building included the Taconic and Acadian orogenies (470–440 and 390–360 million years ago, respectively) and these too may have generated flow systems. But the fluid flow associated with these earlier phases represented only a small harbinger of the massive flow systems created by the subsequent uplift of the Appalachian and Ouachita Mountains and adjacent forelands. (From ref. 9.)

mical and thermal alteration observed throughout the midcontinent. Fracture networks are more likely to have affected regional and local permeability of the Palaeozoic section hosting the ores than to have provided a driving force for regional brine migration.

We prefer a picture involving groundwater circulation at the scale of entire basins, in which the deepest, saline pore water would be driven away from the newly formed mountain belt, updip towards the shallow edge of the foreland platform in its wake, by the raised topography alone⁹ (see figure). A similar gravity-driven flow model may explain patterns of oil migration within the Alberta and Illinois Basins, and fluid flow in the Arkoma Basin.

Regional flow systems driven by topographic relief appear to place many of the pieces in the puzzle of MVT ore genesis in their correct positions. For example, deep flow rates of $1-10 \text{ m yr}^{-1}$ can be sustained for millions of years, although the rates will wane after hundreds of thousands of years as erosion wears down the elevated landscape. Because of the elevated groundwater flow rates, regional heat flow becomes ele-

Large-scale patterns of petroleum accumulation, cementation and diagenesis associated with basin evolution also fit well. The application of such quantitative models are, nevertheless, limited by a lack of quantitative information on the timing and duration of the flows. It is here that isotopic dating studies help to constrain the models of fluid flow. If the isotopic studies could also constrain the durations of the flow systems, an even more fundamental understanding would be possible on the relationships between ore deposits, fluid flow and mountain building in continental terrains. □

Dimitri A. Sverjensky and Grant Garven are in the Department of Earth and Planetary Sciences, The Johns Hopkins University, Baltimore, Maryland 21218, USA.

1. Brannon, J. C., Podosek, F. A. & McLimans, R. K. *Nature* **356**, 509–511 (1992).
2. Nakai, S. et al. *Nature* **346**, 354–357 (1990).
3. Sverjensky, D. A. *Rev. Earth planet. Sci.* **14**, 177–199 (1986).
4. Bethke, C. M. & Marshak, S. A. *Rev. Earth planet. Sci.* **18**, 287–315 (1990).
5. Oliver, J. *Geology* **14**, 99–102 (1986).
6. Ge, S. & Garven, G. *J. geophys. Res.* (in the press).
7. Deming, D. *Geology* **20**, 83–86 (1992).
8. Clendenin, C. W. & Duane, M. J. *Geology* **18**, 116–119 (1990).
9. Garven, G. & Freeze, R. A. *Am. J. Sci.* **284**, 1085–1124 (1984).

Exploding comets

LAST week Daedalus proposed the ultimate rocket fuel. He planned to irradiate solid hydrogen with ultraviolet light in an electric field, to dislodge and displace its constituent electrons and protons. The resulting intensely energetic matrix of monatomic hydrogen, ions and dislocations might easily store twenty times as much energy as TNT.

Daedalus now reckons that nature may have got there first. Far enough from the Sun, in fact at ten times the distance of Pluto's orbit, space becomes cold enough to freeze hydrogen. The tenuous interstellar gas is mainly hydrogen; in those cryogenic regions it must freeze onto any solid surface that offers itself. So, says Daedalus, a comet wandering through the depths of space must over the ages accumulate an outer frosting of solid hydrogen. At the same time it is feebly illuminated by the Sun and stars. By the photoelectric effect, their visible and ultraviolet light will slowly establish an electric field across the frozen, insulating material. Their far ultraviolet, unshielded by any atmosphere, will gradually scramble the protons and electrons of the solid hydrogen into ever more energetic positions. Ultimately, the reacted, radiation-damaged hydrogen surface will reach its maximum possible content of energy. By then it will be buried under new layers of accumulating hydrogen, which will store energy in their turn.

On its long interstellar sojourn, therefore, every comet slowly turns into a king-sized bomb. When it begins to fall towards the Sun, it starts to warm up. At ten times Pluto's orbital radius, sunlight becomes intense enough to melt hydrogen. The newly mobile atoms and ions will recombine in one enormous explosion. The comet will lose its hydrogen overcoat, and will receive a mighty kick in some unpredictable direction.

This, says Daedalus, explains the apparently endless supply of new comets. They are simply old comets blown off course, and appearing along unfamiliar and uncatalogued orbits. A cometary explosion will be sudden, brief, and unpredictable in space and time. Being largely the recombination of monatomic hydrogen, it should radiate mainly in the ultraviolet. Even with his telescope pointing the right way, an astronomer would probably dismiss it as a glitch in his detector electronics. But Daedalus advocates a serious search for such cometary blasts. If a comet later appears on an orbit emanating from that position, his theory will be confirmed. The emission spectrum of the explosion would sensitively reveal the comet's composition, and might serve as a fingerprint to classify it by.

David Jones

Interpreting DNA fingerprints

SIR — Brookfield¹, commenting on the statistical interpretation of forensic DNA typing results, suggested that surveys of diverse human populations using single-locus probes should have revealed patterns of subpopulations characterized by high frequencies of VNTR (variable number of tandem repeat) alleles that are rare in the population as a whole, if such population substructure existed. In fact, the available survey data, generated by the FBI laboratory² and by Kidd *et al.*³, to the extent that they allow subpopulations to be identified and analysed at all, tend to support the hypothesis of population substructure.

The FBI data for caucasians and blacks, although drawn from many US sources, are collected in such a way that most geographical and ethnic distinctions are lost. Even so, for two of six loci surveyed among caucasians, and five of six surveyed among blacks, significant excesses of individuals with single-bin patterns were observed over the frequencies expected if bins could be equated with alleles and the populations were in Hardy-Weinberg equilibrium. The meaning of this excess is unclear because of technical limitations of the assay used to collect the data: closely spaced alleles are unlikely to be resolved, and small alleles may not be detected at all². Nevertheless, the data cannot be taken as strong or unambiguous support for the notion of population homogeneity.

Kidd *et al.*³, in a survey of VNTR allele frequencies among Amerindian populations, discovered several instances of near-fixation of a single VNTR allele in a population, and numerous instances of significant allele frequency differences between populations. These results too can be rationalized: the populations surveyed could represent extreme instances of isolation and small population size. Once again the data do not straightforwardly support the notion that large urban populations are homogenous.

PETER D'EUSTACHIO

Department of Biochemistry,
New York University Medical Center,
New York,
New York 10016, USA

1. Brookfield, J. *Nature* **355**, 207–208 (1992).
2. Budowle, B. *et al. Am. J. hum. Genet.* **48**, 841–855 (1991).
3. Kidd, J. R. *et al. Hum. Biol.* **63**, 775–794 (1991).

BROOKFIELD REPLIES — I agree that it would be beneficial to have further data concerning the degree of population substructure for VNTR loci, and I welcome recent moves to obtain them¹. The important issue is whether we should prevent the legal use of DNA profile information unless and until we have these

data. It is my view that no sufficient case has been made that we should. It does not follow from a legal presupposition of innocence that justice will always be served by ignoring incriminating evidence against defendants.

It is inaccurate to present the debate as being between alternative hypotheses of some substructure and no substructure. Everyone accepts, on both empirical and theoretical population-genetic grounds, that there will be at least a small amount of substructure within 'racial' groups. The important question is whether the degree of substructure will be enough to render seriously inadequate probability calculations assuming none. Calculations reveal that quite large degrees of substructure can leave these calculations fairly accurate. Furthermore, the evidence suggests that the degree of substructure is not large². Thus, except for the special case when both suspect and the truly guilty person are known, *a priori*, to belong to the same inbred racial group, an assumption of no substructure will give probability estimates that are approximately correct.

JOHN BROOKFIELD

Department of Genetics,
University of Nottingham,
Queens Medical Centre,
Nottingham NG7 2UH, UK

1. Anderson, C. *Nature* **355**, 663 (1992).
2. Risch, N. J. & Devlin, B. *Science* **255**, 717–720 (1992).

Right thinking

SIR — Victor Smetacek (*Nature* **355**, 118; 1992) argues convincingly that the ability to write in mirror script reflects latent left-handedness. But I take issue with his extension of these arguments that the right-to-left format of Etruscan and Archaic Roman scripts was essentially conceived in the hemisphere of the brain naturally associated with output via the left hand.

A far simpler explanation is that those scripts, together with Semitic scripts, had their origins among right-handed stonemasons. I am informed by my right-handed friends that a hammer and chisel are naturally picked up in the right and left hands, respectively, so that asked to inscribe a caption on stone, they would naturally move from right to left. Therefore the direction of the script appears likely to have had its origins exclusively in the medium (stone versus parchment, papyrus, paper and so on) on which the script was originally formulated.

DALE SANDERS

Department of Biology,
University of York,
York YO1 5DD, UK

Pulsar emissions

SIR — The recent discovery of pulsed hard X-ray emission from PSR1509–58 (G. Fishman, personal communication) significantly adds to the list of known radio pulsars with energetic emissions. We propose for theoretical reasons an 'activity index' given by B/P , which is proportional to the electrostatic charge on the neutron star (B is the polar surface magnetic field and P the rotation period). The known pulsars with high energy emissions also have the highest value for this index. This index is also inversely proportional to the characteristic age of a pulsar, which suggests that a similarly active pulsar will be found when the debris from supernova 1987A clears sufficiently.

A plausible mechanism for producing hard photons from a radio pulsar is a pair-production cascade¹, given that the hard photons are involved in the cascade. But magnetic opacity would prevent escape of hard photons if emitted near the surface². Phenomenological models^{3,4} have cascading in vacuum gaps high in the pulsar magnetosphere. There is sound theoretical justification for such gaps^{4,5}, and recent calculations show that cascading can occur not only high in the magnetosphere but can also produce dense bunches low in the magnetosphere, which could account for the coherent radio emission⁶.

The accelerating agent in these models is basically the net charge on the system, which scales as the net charge on the neutron star. The net charge is proportional to B/P (refs 5,6). Doubling the charge moves the same potential out a factor of two, decreasing the magnetic field by a factor of eight. This ratio is also proportional to $(P/P)^{1/2}$, where $P/2P$ is the standard pulsar characteristic age.

The table shows the pulsars with large activity indices (small ages); it can be seen that three of the top five have high-energy emissions whereas the other two are quite distant. A strongly magne-

PULSARS ORDERED ACCORDING TO ACTIVITY INDEX

PSR	$(\dot{P}/P)^{1/2}$	B/P (10^{12} gauss s^{-1})	Distance (kpc)
0531+21 (Crab)	3.55×10^{-6}	114	2.0
1509–58	3.20	102	6.67
0540–69 (LMC)	3.08	99	55
1338–62	1.21	39	12.9
0833–45 (Vela)	1.18	38	0.5
1758–24	1.01	32	4.35
1800–21	1.00	32	5.27
1259–63	9.96×10^{-7}	32	2.34
1706–44	9.49	30	1.43
1853+01	8.82	28	3.30
1046–58	8.80	28	2.59
1737–30	8.76	28	3.08
1823–13	8.61	28	5.51
1727–33	7.75	25	4.13
1930+22	6.31	20	6.58

Data from A. G. Lyne, R. N. Manchester and J. H. Taylor (personal communication)

Women on the verge

Margaret W. Rossiter

To the Ends of the Earth: Women's Search for Education in Medicine. By Thomas Neville Bonner. *Harvard University Press: 1992. Pp.232. \$34.95, £27.95.*

TODAY there are women physicians nearly everywhere. But it was not always so. Thomas Bonner now provides a fresh examination of the different strands of their long and intense struggle for medical training. Between 1850 and 1908, women were admitted to medical schools in most Western countries as the result of two developments — the creation of separate medical schools for women (as happened in the United States, the United Kingdom and Russia), and the later but direct admission of women to full coeducation at traditional institutions elsewhere.

When, in the nineteenth century, the first women applicants to medical schools were rebuffed, their supporters could retaliate by starting their own fledgling medical schools for women. This provided training and even degrees — a kind of entering wedge — for the first women practitioners. It even provided jobs eventually for a few female medical professors. Initially, the quality of this education was not an issue, for that of the men-only institutions was not much better. Yet some of the early pioneers, wealthier or more ambitious than the others, felt that the level of medical training was second-rate, and so they travelled abroad in search of more rigorous schooling.

One of the first institutions in the world to admit foreign women, the University of Zurich, trained thousands of female physicians between 1876 and 1914. Many, if not most, were Russians, intent on educating themselves so that they would be able to help modernize their backward nation. They received considerable criticism from townspeople, although faculty members (who included many professors too liberal for the German universities of the time), university rectors and cantonal officials defended the women and their often — for the time — shockingly unladylike behaviour, which included wearing short hair, smoking, going about unchaperoned and becoming involved in radical politics. As the number of women seeking foreign training grew, because of intermittent bans on medical training at home by vacillating tsars, and pogroms in Russia and Eastern Europe, the number of women at other welcoming medical schools, such as those in Bern, Paris and Geneva, also increased.

Bonner devotes most of the second half of his book to the story of what

happened when the women returned home and tried to open their local institutions to other women. Their success varied widely. Because medical training in Europe was controlled by state-run universities and they, in turn, were ruled by government ministries, many individuals became involved in the momentous decision to admit women. Thus applicants and women's organizations could send their petitions (one had 55,000 signatories) to faculty members, rectors, officials of the ministry of education, or even the Reichstag, and any of these authorities could oppose the petitions. Even then, admission could be sabotaged by the discourteous conduct of a faculty or student.

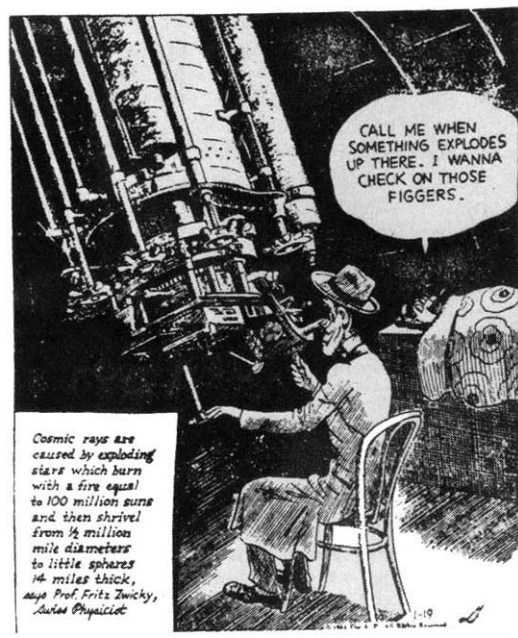
Bonner is not totally clear on why men's reactions to the admission of women varied as they did, ranging from the enlightenment of one particular rector at the University of Bern to the hysterical and stubborn aversion of other individuals. In 1900, however, Austrian universities began to accept women medical students, and so, in 1908, did even those of Prussia. Bonner claims that once the battle for coeducation was won, the doors, in some cases propped

open by the medical needs of European wars, stayed open (except for a time in Hitler's Germany).

From this perspective, the United States and the United Kingdom, countries usually praised in the history of women's medical education for their early 'firsts' such as Elizabeth Blackwell, Sophia Jex-Blake and Elizabeth Garrett Anderson, are here seen as laggards. The reliance of these countries into the early 1990s on separate and private institutions, which was both possible and necessary where higher education was outside the rule of government officials, meant that they were powerless to prevent the prevailing discrimination. The provision of such ready alternatives at home merely helped to perpetuate the discrimination. But one could also argue that these countries closed their separate colleges too soon, for the number of women enrolling at the coeducational schools was not large, and discriminatory quotas persisted into the 1960s. (They were ended in the United States as late as 1970 with the lawsuit brought by the Women's Equity Action League against what were, by then, "federal contractors".)

One suspects that Bonner's upbeat tale may be a bit too optimistic, for once women had won admission into and graduation from educational institutions, their full participation in medicine was often then limited by exclusion at the next level up, for example, from internships and medical-society memberships. Bonner mentions this only

Be Scientific with OL' DOC DABBLE.



Ol' Doc Dabble — the earliest publication of the neutron star model for supernovae in January 1934. The model was first suggested by Fritz Zwicky and Walter Baade just two years after the discovery of the neutron. This picture is taken from a chapter on supernovae and stellar catastrophe by Robert P. Kirshner that appears in *Understanding Catastrophe*, a book originating from the 1990 Darwin College Lectures. Other contributors include Walter Alvarez on dinosaur extinction; Martin Rudwick on Darwin and catastrophe; Claudio Vita Finzi on earthquakes; Nicholas Cook on storms and cyclones; Peter Garnsey on famine; and Roy Porter on changing attitudes to disease. Published by Cambridge University Press, price £17.95, \$29.95.

briefly. But in general he tells the complex multinational tale clearly and well, although at times he is repetitious and didactic. From his research in Switzerland and the United States he has unearthed much relevant new material on the early women pioneers, including some letters that they wrote back to home and to each other. And it is also clear that these pioneers did not remain totally obscure: in fact, their presence was often such a *cause célèbre* as to

generate both derisive and supportive cartoons. Bonner's book is the best one-volume survey of this complex subject, and deserves wide use in the increasing number of courses on women's history, women and science, and the history of medicine. □

Margaret W. Rossiter is in the Department of Science and Technology Studies, Cornell University, 632 Clark Hall, Ithaca, New York 14853, USA.

Worldly wise

Peter J. Smith

Encyclopedia of Earth System Science. Volumes 1–4. Editor-in-chief William A. Nierenberg. *Academic*: 1992. \$950, £573.

WHAT is this new 'Earth system science' and, in particular, how does it differ from plain old Earth science? The answer to the second part of this question, which also happens to be the answer to the first, is that there is no difference at all in substance but only in emphasis, and that the alleged novelty of the emphasis is more imaginary than real.

"In the past few decades", claims Robert M. White, president of the US National Academy of Engineering, in the foreword to this huge encyclopaedia, "humanity has come to understand that all parts of the earth system, the oceans, the atmosphere, the solid earth, the biosphere, and the near space surrounding it, make up a continuously interacting whole." I doubt it. The vast mass of humanity has never given such matters even a passing thought; and if by 'humanity' White simply means those involved, however peripherally, in and with the Earth and environmental sciences, they were aware of the interconnectedness of Earth processes long, long ago. The literature is replete with examples of solid-Earth-oceanic-atmospheric-biospheric interactions, ranging from the outgassing from the solid Earth of an atmosphere and oceans subsequently modified by chemical, biological and human activity to the way in which processes deep in the Earth influence inhabitants of seismic and volcanic zones.

The problem, as it always has been, is not lack of awareness of terrestrial interactions but the ability to understand them in scientific detail. And therein perhaps lies a tale, for greater scientific understanding needs money, and lots of it. There may be little new in Earth system science apart from the name, but cynics would argue, applying the Windscale-Sellafield principle, that re-

generating is precisely the point. Now that politicians have become interested in Earth-environmental matters, which by their very nature involve complex natural interactions, there may be considerable financial advantage to be had in pretending that poised to solve the problems is not some fuddy-duddy-sounding science with 200-year-old roots but a glossy new concept dressed in high-tech jargon.

In this context, the *Encyclopedia of Earth System Science* can be seen as part of the talking-up, or softening-up, process. It may be simply an encyclopaedia of the Earth sciences under another name and it is certainly not, as White appears mistakenly to believe, any more interdisciplinary than would be any other encyclopaedia of the Earth sciences published in 1992. It neither records a quantum leap in interdisciplinary understanding nor, obviously, is able to take such understanding beyond what the current state of knowledge will allow. What is true, however, is that it is the first *multidisciplinary* encyclopaedia of the Earth sciences to appear; and there may be some merit in that if it encourages interdisciplinary studies in the future.

What we have, then, are four tomes comprising more than 2,600 pages of text and containing 230 articles with an average length of about 12 (large) pages each, although there is considerable variation about the mean. Each article includes, in order, a list of the sections therein, a general abstract-style introduction, a more specific introduction to the subject, a text divided into sections, a modest bibliography emphasizing books and reviews, and a short glossary. The final volume also contains a list of contributors, a 13-page 'relational index' comprising general headings under each of which is a list of relevant articles (very useful), and a detailed 171-page subject index (even more useful).

The subjects covered range from acid rain, acoustical oceanography, Antarctica and atmospheric tides, through magma chamber dynamics, marine pollution, the Messinian salinity crisis and microplates, to tsunamis, volcanic tremor, wood moisture and the environment, and zonal and meridional winds in the Earth's atmosphere. As this short list

covers just 5.2 per cent of the articles, it can do no more than give a flavour of the contents, although I hope it is sufficient to illustrate the huge range of topics included.

What I have read and skimmed — which is about as much as one can reasonably do in an average working week — has been generally well written and, indeed, so uniformly well written that I cannot but conclude that, unusually in these days, the volumes have actually been edited and not merely compiled. I could have wished that the articles had been equally uniformly comprehensible, although perhaps that was too much to hope for given natural variations in subject complexity. Thus whereas many of the articles will be easily understood by the "policymakers and administrators" that Nierenberg hopes to include among his audience, many will not, and some (for example, that on the origin of the geomagnetic field) will be accessible only to those with PhDs in, and several years' experience of, advanced mathematics (which does not include me).

One cannot but be impressed by the scale and quality of the enterprise, although the former must be put in perspective. In spite of its great length, this new encyclopaedia is actually less comprehensive within individual fields (such as the solid Earth) than are some recent, more specific volumes (such as *The Encyclopedia of Solid Earth Geophysics*, Van Nostrand Reinhold, 1989). More importantly, however, what influence is it likely to have? No individual will be able to afford it; it will be beyond the resources of any UK public library, given the recently publicized crisis in that system; most university libraries are in an equally parlous position; and it is not the sort of thing that could usefully be borrowed for a couple of weeks from the National Lending Library. So who, in the United Kingdom at least, will ever see it? □

Peter J. Smith is in the Department of Earth Sciences, The Open University, Milton Keynes MK7 6AA, UK.

Addendum

Apoptosis: The Molecular Basis of Cell Death (Cold Spring Harbor Laboratory Press), reviewed in the issue of 5 March, is edited by Frederick O. Cope, in addition to David Tomei.

Spring Books

Appearing in next week's issue is *Nature's* Spring Books supplement. Featured, among others, are Huw Price on the enigma of Paul Davies; Richard Davenport-Hines on miasmas and disease; Colin Renfrew on the Atlantis legend; Walter Gratzer on Otto Warburg; David L. Hull on sexual selection; David M. Ritson on Edward Teller; Ed Regis on Caltech; and Ian Stewart on Buckminster Fuller's "scenario for the future of humanity".

Mutual gains in a cruel world

Marian Stamp Dawkins

The Covenant of the Wild: Why Animals Choose Domestication. By Stephen Budiansky. *William Morrow*: 1992. Pp. 190. \$18.

THIS book puzzled me. On the one hand, I found it refreshing to read Budiansky's coherently argued but unfashionable views on fashionable topics such as the ethics of hunting and farming. On the other, I felt let down because, having said that many people have the wrong ideas about farms and domestication, he does not follow this

the human where the honey is and the human opens the nest, giving the bird some of the spoils. At other times, the association has become more intimate and the animal is now no longer wild, as with domesticated dogs, although there is still mutual benefit. Dogs gain food, protection and companionship, while humans gain a great deal from the presence of dogs. Even farm animals such as sheep and horses, the author argues, are symbionts, gaining from the protection of humans and living in far greater numbers around the world than they ever would if left to their own devices. He also criticizes the romantic view that some animal-rights supporters have of life in the wild. People are easily shocked by the horror stories of the laboratory, he says, but are ignorant of the greater horror stories of nature, such as

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Who's kidding whom? An eleventh-century illustration of animal domestication.

through to saying what he thinks the right ideas might be. I was tempted to keep asking, "So what?"

Budiansky's thesis is that human relationships with other species are not so much the 'exploitation' they are often made out to be, but arise from quite natural coevolutionary processes like those between other pairs of species. For example, acacia trees grow special shelters for ants. In return, the ants defend the trees against herbivores. The association is not one of the ants 'exploiting' the trees or the trees 'duping' the ants, but one of mutual cooperation. Similarly, birds flocking together are all selfishly making use of the fact that 20 pairs of eyes are better than one. All individuals gain from this behaviour, even though they could be described as mutually parasitic on each other.

In Budiansky's view, many human associations with animals are mutualistic in nature, with both partners, not just the human masters, gaining in the process. Sometimes, as with honey guides (birds that eat honey but cannot open a bees' nest without assistance from a larger animal), human and wild animal both have obvious gains: the bird shows

the death and killing that are regular features of the lives of wild animals.

So, because Budiansky argued that many of the ways in which humans behave towards other animals are no different in principle from the ways in which animals behave towards each other, I was expecting him to arrive at some sort of conclusion about the moral issues surrounding human treatment of animals. Was he about to say that as so many things (even killing) are 'natural', we should not worry about them from an ethical point of view? Or was he going to say that we still should, despite the far worse things that go on in nature? When he said neither, I felt as though I had been left in the middle of nowhere after the last bus had gone. But, on reflection, perhaps it wasn't a bus that I missed, but a bandwagon. Perhaps the strength of the book is that it raises issues and makes points and then doesn't tell one what conclusions to draw. It is readable and thought-provoking, and doesn't pretend to be anything more. □

Marian Stamp Dawkins is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

Channel cheer

Charles F. Stevens

Electrogenic Ion Pumps By Peter Läuger. *Sinauer/W. H. Freeman*: 1991. Pp. 313. \$44.95, £34.95.

ALL cells live or die by their ion pumps, those marvellous molecular machines that use energy to move critical ions — sodium, potassium, calcium, chloride and hydrogen — to where the cell needs them. Peter Läuger, who died a year-and-a-half ago in a tragic accident, has left us a modern, readable and scholarly review of this important class of molecules.

The pump familiar to all biologists, the sodium-potassium ATPase that keeps potassium on the inside and moves sodium to the outside of all animal cells, is but one member of a larger family of pumps that use ATP hydrolysis to transport ions. But nature has invented other families of pumps that use other sources of energy: light, phosphate hydrolysis, decarboxylation and the transfer of electrons in oxidation-reduction reactions. One of the strengths of this monograph is that all of these pump classes are treated in the context of the general principles of coupling energy expenditure to ion transport.

Another strength is that information from the various approaches to pumps is well integrated. Results from biochemical, ultrastructural, physiological and molecular biological studies are all combined with the sophisticated theoretical perspective that was Läuger's trademark. Although an appropriate amount of theory is here, this is not primarily a theoretical treatment of pump function. Indeed, it is the experimental approaches to pumps that predominate in the discussion. Pump molecular biology has been moving rapidly, and Läuger might well have included more from this area were he able to write the book now.

Electrogenic Ion Pumps is divided into two parts, the first a discussion of common principles and the second a consideration of specific pumps. The first section begins with an overview of known ion pumps and what they are good for. Most of this part is devoted to the basic principles of pump function and a survey of the way they are studied. This is where the theory is.

The second section treats six specific pumps, one chapter for each. Here bacteriorhodopsin, the *Neurospora* proton pump, the sodium-potassium ATPase, the sarcoplasmic reticulum calcium pump, proton-pumping ATPases and cytochrome oxidase are discussed systematically and in detail. Each chapter

summarizes our current knowledge of structure and function in addition to treating other special topics, such as regulation or comparison with other relevant molecules.

Although the book is clearly organized and provides a coherent treatment of the subject, one need not read it straight through. The two parts can, depending on the level of sophistication the reader brings to the material, be read independently and out of order, and each of the chapters treating a specific pump constitutes a brief, free-standing review of that pump.

From know-how to nowhere

Henry Petroski

Flying Buttresses, Entropy, and O-Rings: The World of an Engineer. By James L. Adams. *Harvard University Press*: 1992. Pp. 264. \$24.95, £19.95 (hbk); £10.95 (pbk, UK and Europe only).

THIS book is a wide-ranging discussion of engineering from one engineer's perspective. It is intended primarily for general readers, for lay people who interact in their job or life with engineers, and for those thinking about studying engineering or seeking to add a technological component to their liberal education. Adams explains in his introduction that, as an engineer, he has been "often embarrassed by the relatively small number of readable general books on engineering, as opposed to the large number on the physical sciences and mathematics", and it is, in part, this paucity that led him to write his book.

What makes a book on engineering or any other subject readable is, of course, a highly subjective matter, and the problem of objectivity is compounded for an engineer reading a book intended mainly for nonengineers. Adams anticipates this, and in his introduction speaks to the professional, acknowledging that no two engineers are likely to have the same viewpoint and perspective on the "complex and multifaceted activity" in which they are engaged. Indeed, Adams hopes that engineers who experience too much of a mismatch between his perspective and theirs will write their own books to "help address the biases" in his. With such an apologetic and self-conscious introduction, Adams virtually throws down the gauntlet before the reviewer.

It is usually easy to nit-pick about what is wrong with a book in one's field, but that would serve little purpose here unless the points of disagreement were

This lucid summary of ion pumps will be widely appreciated by those in the transport field, and it provides a fitting memorial for one of the field's intellectual leaders and innovators. □

Charles F. Stevens is in the Molecular Neurobiology Laboratory, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

■ The second edition of *Ionic Channels of Excitable Membranes* by Bertil Hille has also recently been published by Sinauer/W. H. Freeman, price \$46.95, £37.95. For a review see *Nature* **314**, 388; 1985.

materially to change the impression of engineering that the nonspecialist will take from this book. Adams does an admirable job of setting engineering in a broad historical perspective, pointing out that modern technology has very ancient roots. He also paints a very fair picture of the complexity of the endeavour, and demonstrates the origins of engineering problems and the variety of modes in which engineers work to attack them. The centrality of design and invention is discussed, as are the roles of mathematics, science, and research and development in leading up to the moments of truth embodied in production, manufacturing, assembly, testing and use. The importance of economic factors in engineering is discussed, and the implications for technology and society of regulation and participatory democracy are explored. In short, Adams covers in 250 pages or so just about as much as can be expected of an endeavour that touches and is touched by just about every aspect of civilization and culture.

The general reader is thus likely to get a good idea of what engineering is and what engineers do. How readable the book is will no doubt depend on the receptivity of the reader and the extent to which the book succeeds in drawing more combative readers into its flow. In this regard, some of Adams's seeming asides may prove to be obstacles. For example, in discussing the use of the computer to solve differential equations formerly solved by hand, Adams tells us that this technological development has left engineers time for "more important kinds of thinking", but he does not say what kinds of thinking are more important for an engineer. In discussing what students should be expected to do in a design project, he describes the necessity of finding information and then rather glibly admits that the first thing he himself does when seeking information is to ask someone, because he likes people better than libraries. No matter how many caveats Adams has put in his introduction about this being his personal point of view, such a remark (and its implied advice to students) will all too easily reinforce the caricature of the engineer as anti-intellectual, not to mention raising the eyebrows of any humanists who might read the book. □

Henry Petroski is in the Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina 27706, USA.

■ Of related interest is the second edition of *The Maze of Ingenuity: Ideas and Idealism in the Development of Technology* by Arnold Pacey. The book now contains several new chapters. Published by MIT Press, price \$25, £22.50 (hbk), \$12.95, £11.75 (pbk).



In suspense — Brooklyn Bridge under construction in 1878. Taken from *The Engineer in America* edited by Terry S. Reynolds, a historical anthology of essays originally published in the journal *Technology and Culture*. The book traces the rise of the profession over the past two centuries. Published by The University of Chicago Press, \$44.95 (hbk), \$19.95 (pbk).

The end of cold dark matter?

M. Davis, G. Efstathiou, C. S. Frenk & S. D. M. White

The successful cold dark matter (CDM) theory for the formation of structure in the Universe has suffered recent setbacks from observational evidence suggesting that there is more large-scale structure than it can explain. This may force a fundamental revision or even abandonment of the theory, or may simply reflect a modulation of the galaxy distribution by processes associated with galaxy formation. Better understanding of galaxy formation is needed before the demise of CDM is declared.

How did structure in the Universe form? This question has puzzled mankind for centuries, but in the past decade some cosmologists have felt that they were close to providing an answer. What has become known as the cold dark matter (CDM) theory is an elegant construct which links many aspects of the structure we see today to physical processes which took place when the Universe was only 10^{-35} s old. Recently, observations have been reported that seem to conflict with this model (see, for example, ref. 1). Should we continue to be impressed by the successes of the CDM model, by the simplicity of its underlying assumptions and by its link to the physics of the early Universe, or do these new data require us to reject it and to seek a different explanation for observed structure? If the situation is as yet unclear, what key observations could test the model definitively?

Any theory for the formation of cosmic structure must specify two chief ingredients: the amount and nature of the material that fills the Universe, and the properties of the seed density fluctuations from which galaxies and clusters developed. There is now overwhelming evidence that most of the mass in the Universe is invisible, detected only by its gravitational effects on stars and galaxies. The total amount is poorly constrained by observations and could, perhaps, be sufficient to halt the cosmic expansion. Although this dark matter might be nothing stranger than planets or dead stars², particle physicists have suggested more exotic possibilities such as gravitinos, photinos, massive neutrinos and axions³. Of these weakly interacting particles, known collectively as nonbaryonic dark matter, only the neutrino is known to exist and even this particle may not have sufficient mass to dominate the Universe. Apart from hypothetical nonbaryonic dark matter, particle physics has also suggested ways of generating the seed fluctuations required for galaxy formation, all related to phase transitions that may occur as the Universe expands and cools. The CDM model is based on a particular class of nonbaryonic dark matter and on a simple type of seed fluctuation for which many aspects of structure formation can be calculated with some confidence. In this model we would expect the large-scale structure of the present Universe to bear a strong imprint of the earliest phase of its evolution.

Cosmologists have worried for many years about the 'initial conditions' of the Big Bang. Why is our Universe so smooth? Why has it lasted so long without either collapsing again or going into free expansion? What is the origin of the seed fluctuations required to make galaxies, stars and people? It seems implausible that the physics of the first few instants should favour scales relevant to structures in the present Universe, and so it has often been argued that the Universe emerged from the Big Bang without any characteristic scales imprinted on it. Its radius of curvature must then be much larger than its current observable size, and the amplitude of fluctuations in curvature must be independent of their wavelength^{4,5}. These notions received a tremendous boost in the early 1980s from the idea of an inflationary phase in the initial stages of the Big Bang⁶⁻⁹. According to this idea, an early phase transition subjected the Universe to a period of exponential expansion during which its size increased enormously. As a result, its present curvature is immeasurably small and the seed fluctuations for galaxy forma-

tion could have originated from quantum fluctuations that were inflated to macroscopic scale. Except in circumstances that appear contrived, the fluctuations would indeed contain no characteristic scales; in technical terms, irregularities in the spatial curvature are predicted to be a gaussian random field with a scale-invariant spectrum⁹⁻¹². For the first time cosmologists had a set of initial conditions stemming directly from fundamental, even if speculative, physics.

A spatially flat universe is most simply achieved if the mean mass density has the critical value needed to halt the expansion. This is a bold prediction, as it directly contradicts two well entrenched observations. Dynamical analyses of galaxy clusters consistently measure a mass per galaxy which, if extrapolated to the Universe as a whole, provides no more than 10–20% of the critical density^{13,14}. In addition, the standard theory of Big Bang nucleosynthesis explains the observed abundance of light elements (H, He and Li) only if the present density of ordinary (baryonic) matter is less than 10% of the critical value¹⁵. A theoretical argument against a purely baryonic Universe, flat or otherwise, is that the seed fluctuations expected from inflation cannot then account for observed structure because all fluctuations on scales smaller than galaxy clusters are erased at early times¹⁶. Zel'dovich had therefore argued in the early 1970s that galaxies must have formed as fragments of collapsing clusters¹⁷. There is now evidence, however, that clusters are younger than the galaxies within them, and Zel'dovich's scheme also predicts ripples in the microwave background radiation¹⁸ that are ruled out by recent observational data. As a result this 'top-down' view of cosmogony has fallen out of favour.

In 1980 a Soviet experiment claimed to measure a mass for the electron neutrino sufficient for such particles to dominate the density of the Universe¹⁹. The announcement was greeted with excitement, because the difficulties mentioned above might be circumvented if the dark matter were nonbaryonic. The dynamical estimate of the mean mass density derived from galaxy clusters depends on an unjustified assumption that the mass per galaxy in clusters is representative of the Universe as a whole; the nucleosynthesis constraints apply only to baryonic matter that participates in the synthesis of chemical elements, not to nonbaryonic dark matter; some types of nonbaryonic dark matter do not require clusters to form before galaxies; and a gravitationally dominant nonbaryonic component can promote structure formation while leaving only a weak imprint on the microwave background radiation (see Fig. 1). But the neutrino mass measurement was not confirmed by later experiments, and interest in neutrino universes waned as other problems emerged. For example, it became clear that the rapid thermal motions of such 'hot dark matter' would erase all small-scale structure at early times. ('Hot' in this context refers to light particles, such as neutrinos with masses ≤ 100 eV, that have high enough thermal speeds to damp out structure smaller than galaxy clusters.) As in a baryonic model, a neutrino-dominated universe with the 'natural' fluctuations predicted by inflation would lead to a cosmogony in which galaxies formed by the fragmentation of larger structures. As Fig. 2 shows, such a universe appears very unlike our own²⁰. For other types of

weakly interacting particles ('cold dark matter'), slower thermal motions are predicted. Galaxy-sized fluctuations can then grow and collapse before clusters, explaining the apparent sequence of cosmogony²¹. A universe of cold dark matter looks considerably more like the real world (Fig. 2), but whereas neutrinos are known to exist, no 'cold dark matter' candidate has yet been detected in a laboratory.

The CDM model

In the standard CDM model, we assume a flat universe in which 5–10% of the critical density is provided by ordinary matter and the rest by the hypothetical weakly interacting cold particles. In addition, the seed fluctuations are assumed to be of the scale-invariant form discussed above. The key assumptions thus fit in nicely with ideas on inflation. In the resulting cosmogony the first objects to form are small clumps of gas and dark matter, similar in mass to dwarf galaxies (see Fig. 3). These coalesce under the action of gravity to make objects the size of our own Galaxy which in turn aggregate into clusters and superclusters of galaxies. This 'bottom-up' scheme for structure formation is a version of the 'hierarchical clustering' model that Peebles and others investigated in the 1970s²². Unlike the original models, however, it does not require contrived seed fluctuations or suffer from the difficulties associated with baryonic universes. Because the growth of large-scale structure in a CDM universe is promoted by gravity, the theory also accords with recent results from the COBE satellite which show that the spectrum of the microwave background fits a black-body curve to extraordinarily high precision²³. This sets tight limits on the amount of energy that could have been injected into the intergalactic medium at recent epochs, and eliminates models in which structure forms nongravitationally, for example from shock waves driven by exploding galaxies²⁴ or by superconducting cosmic strings²⁵. Direct evidence that gravity is the prime agent shaping the observed structure comes from the measured density and velocity fields in the local Universe²⁶. Galaxies do indeed seem to be

moving towards galaxy clusters, as expected if structure is induced by gravity.

Apart from agreeing with the *a priori* prejudices of cosmologists, the CDM model provides a quantitative explanation for many of the fundamental properties of galaxies and clusters. In 1984, Blumenthal *et al.*²⁷ outlined the attractive features of the model, and much work since then, including our own computer simulations²⁸, has supported their initial optimism. Gravitational clustering produces CDM clumps with masses ranging from those of small galaxies to those of rich clusters of galaxies, and it produces them in about the right abundance to match observation. The luminous parts of galaxies form when gas radiates its binding energy and settles to the centre of its CDM halo, spinning up in the process. CDM replaces the *ad hoc* initial conditions that were assumed when this two-component picture of galaxy formation was first developed²⁹ with just those that seem to be needed to explain the masses, sizes, angular momenta, and abundance of galaxies and galaxy clusters. Recently, it has become possible to simulate CDM universes including both dark matter and gas. Although such work is in its infancy, it is already clear that cooling of gas within a collapsing protogalaxy can lead to a disk which resembles a spiral galaxy³⁰, and that clusters of galaxies are expected to have gaseous atmospheres similar in structure to those of real clusters³¹.

Our early work³² and that of Bardeen *et al.*³³ confirmed that galaxy clustering and the associated galaxy motions are inconsistent with the CDM model unless galaxies and dark matter are distributed differently. In principle the relation between the two could be calculated by numerical simulations which follow galaxy formation in detail. In practice this is difficult because of the large computer resources required and, more importantly, because of uncertainties in the physical processes involved. As a result most studies have assumed a simple model in which galaxies form from matter located near high peaks of the initial (low-amplitude) CDM density field. Many aspects of this 'high-peak' model can be calculated analytically³³, and it is easily implemented in computer simulations. It produces a biased 'galaxy' distribution with higher density contrasts than the underlying dark-matter distribution, as required if galaxy clusters are to give misleadingly low estimates of the global mass per galaxy. But it is a model of convenience, and it may correspond poorly to the way that galaxies would actually form in a CDM universe.

Our first attempts to simulate galaxy formation suggested that a bias similar to that in the high-peak model would indeed arise³⁴. More recent simulations^{35,36} extend this and suggest that several biasing mechanisms may be important. The modelling of hydrodynamical processes and star formation remains crude, however, and further work on the relationship between the galaxy and the CDM distributions is badly needed. Studies using the high-peak model have shown that CDM universes can match the observed pattern of galaxy clustering on small scales³², its 'sponge-like' topology³⁷ and the abundance of galaxy clusters³⁸. They even come close to matching the spectacular coherent structures seen in three-dimensional surveys of the local Universe, objects such as the giant void in Bootes^{39,40} and the 'Great Wall'^{41,42}. These structures are sometimes claimed to pose problems for the CDM model, but to our knowledge there has been no convincing demonstration of a quantitative discrepancy so far²⁸.

Discrepancies with the CDM model?

The most serious observational challenge to the CDM theory has come from statistical studies of the galaxy distribution on large scales ($\geq 10 h^{-1} \text{ Mpc}$; here and below we write the Hubble constant as $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$). A recent catalogue of more than 2 million galaxies reveals more large-scale structure than predicted by simple versions of the theory⁴³ (see Fig. 4). Further evidence has come from redshift surveys of optical

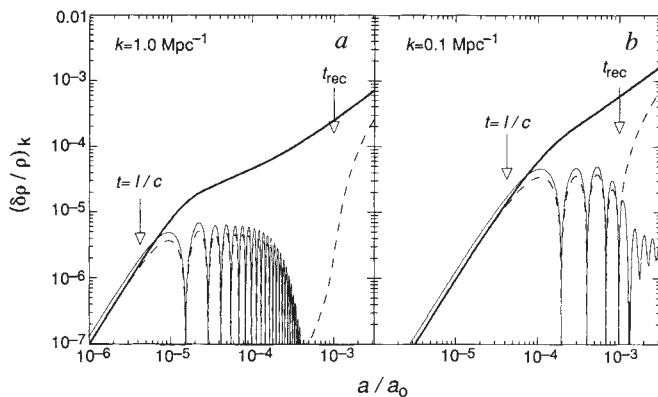


FIG. 1 Evolution of small fluctuations in a critical-density CDM universe with Hubble constant $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The figure shows the amplitudes $(\delta\rho/\rho)_k$ of plane-wave fluctuations of wavelength $\lambda = 2\pi/k$ for (a) $k = 1.0 \text{ Mpc}^{-1}$ (roughly the size of a large galaxy) and (b) $k = 0.1 \text{ Mpc}^{-1}$ (about the size of a cluster of galaxies). The abscissa shows the scale factor a of the universe (which is related to time by $a \propto t^{2/3}$) divided by its present value (a_0). Thick solid lines show the amplitude of fluctuations in the cold dark matter. Dashed lines show fluctuations in the baryons (which in these calculations account for 5% of the critical density), thin solid lines show fluctuations in the radiation energy density. Arrows marked $t = l/c$ show the light travel time across single wavelength. Before $t = l/c$, all of the perturbations grow as a power law in time, but after this epoch baryon and photon fluctuations oscillate like sound waves with a near-constant or decaying amplitude. Arrows marked t_{rec} indicate the time when electrons and protons 'recombine' to form hydrogen, and the universe becomes transparent to radiation. As CDM fluctuations can grow between $t = l/c$ and t_{rec} , it is possible to produce large-scale structure while generating small anisotropies in the microwave background radiation that are below current upper limits.

galaxies (J. Loveday, G.E., B. A. Peterson and S. J. Maddox, manuscript in preparation) from infrared galaxies detected with the IRAS satellite^{1,44}, from clusters of galaxies⁴⁵ and from radio galaxies⁴⁶. Although this seems an impressive list, the structures detected have relatively small amplitude, with r.m.s. fluctuations of $\sim 30\%$ in the number density of galaxies on scales of $20h^{-1}$ Mpc (about 2–3 times the amplitude expected in the standard CDM model). Measuring fluctuations at this level is difficult, and small errors or biases in the catalogues can masquerade as large-scale structure. Two currently controversial examples involve the clustering of galaxy clusters and of IRAS galaxies. There is substantial evidence that some component of the strong clustering in Abell's catalogue of galaxy clusters, compiled more than 30 years ago by visual inspection of photographic plates, is actually spurious⁴⁷. A new cluster catalogue constructed by computer analysis of digitized photographic plates is more weakly clustered and closer to the predictions of the CDM model⁴⁸. Large-scale structure in a sample of IRAS galaxies suggested the possible death of CDM¹. But an analysis of a new and substantially larger redshift survey of IRAS galaxies (K. Fisher, M. D., J. P. Huchra, M. Strauss and A. Yahil, manuscript in preparation) indicates less clustering than the original study, although the two surveys may not be inconsistent with each other. The problem here is that statistical uncertainties in measurements of clustering are notoriously difficult to estimate and so it is often hard to judge the significance of any discrepancies.

Thus, at present, the strongest evidence for excess clustering on large scales comes from recent two-dimensional galaxy catalogues. These provide a higher level of statistical significance than galaxy redshift surveys because they sample much larger volumes of space. But they also need careful control of systematic errors because the amplitude of the clustering in two dimensions is reduced by projection onto the sky. As noted above, differences between existing redshift surveys may be due

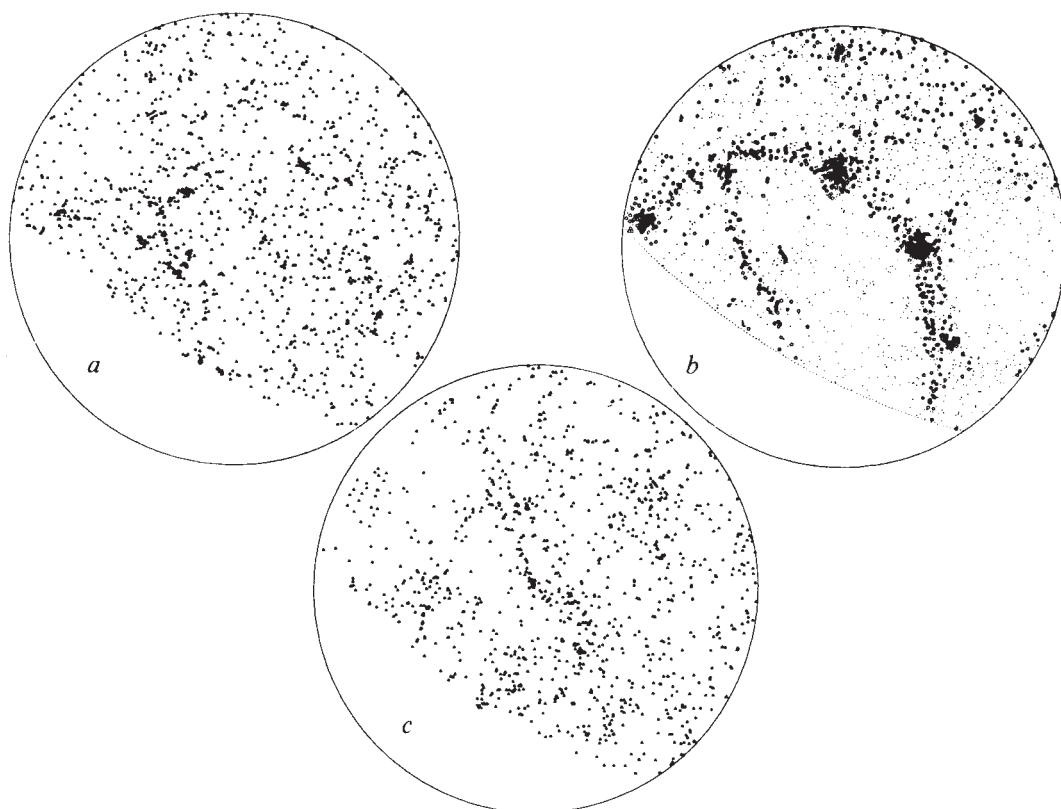
to sampling fluctuations or to small selection biases. Clearly, we would like larger redshift surveys and better checks on the uniformity of the two-dimensional catalogues.

Observational data on the peculiar motions of galaxies (their deviations from a pure Hubble flow) have improved enormously in recent years through large observing programs⁴⁹. As a result it is becoming possible to use the velocity field to map the mass distribution in the local Universe^{50,51}. Some authors⁵² have argued that current observations already imply that these streaming motions are coherent over larger scales than predicted by CDM, but other analyses suggest that the random and systematic uncertainties in these data are too large to rule out the theory^{53,54}. As new data accumulate, this technique should yield important constraints. A particularly strong point in its favour is that velocities are directly induced by the mass distribution and, at least in the standard model, are largely independent of the detailed relationship between the mass and galaxy distributions. It should be borne in mind, however, that even a weak dependence of galaxy properties on environment could confuse the interpretation of peculiar velocity data.

Do we live in a CDM universe?

Why does the presence of large-scale structure in the Universe cause difficulties for the CDM model? A CDM universe does in fact develop a characteristic scale even though its initial conditions are scale-free. This scale arises when CDM takes over from the fireball radiation as the gravitationally dominant constituent of the Universe and is equal to about $10(\Omega h^2)^{-1}$ Mpc, where Ω is the ratio of the present mean density to the critical density. The model predicts that clustering should be relatively weak on scales much larger than this. Hence lowering either the density or the value of the Hubble constant in a CDM universe increases the degree of clustering on large scales. Our numerical simulations required $h \approx 0.5$ to match observations of galaxy clustering on scales $\leq 10 h^{-1}$ Mpc. As this is already at the low

FIG. 2 Equal-area projections of the galaxy distribution on the northern sky and in artificial catalogues made from N -body simulations. *a*, Artificial catalogue in a standard CDM model with biased galaxy formation. *b*, Artificial catalogue in a neutrino-dominated model in which galaxy formation began at a redshift of 2.5; dots represent the neutrino distribution, circles represent the galaxies. *c*, Galaxies in the Center for Astrophysics northern survey. The outer circle represents galactic latitude $+40^\circ$, and the empty regions lie at declinations below 0° .



end of the range allowed by observation⁵⁵, lowering the Hubble constant still further seems an implausible way of obtaining more large-scale structure. Lowering Ω is another possibility, but without an additional ingredient such models are inconsistent both with a spatially flat universe and with present upper limits on fluctuations in the microwave background^{56,57}. These problems can be avoided by appealing to a cosmological constant, because a low-density universe is spatially flat if the cosmological constant takes the value⁵⁸ $\Lambda = 3H_0^2(1 - \Omega)$. With such carefully chosen parameters it is possible to construct a CDM universe that explains large-scale structure⁵⁹, is compatible with inflation and with microwave-background experiments, and is old enough to contain the oldest observed star clusters even for a present expansion rate as high as $H_0 = 80 \text{ km s}^{-1} \text{ Mpc}^{-1}$, the value preferred by some recent measurements^{60,61}. From the point of view of a particle physicist, the value of Λ needed to work these miracles is extraordinarily small, 10^{120} times smaller than its 'natural' value⁶². Such fine tuning seems sufficiently unattractive that most cosmologists regard this solution as a long shot, preferring to think that some unknown symmetry principle requires the cosmological constant to be exactly zero.

Other possible fixes for the CDM model involve decaying particles or departures from the scale-invariant seed fluctuations predicted by simple inflationary models. For example, the predictions for large-scale structure can be boosted if 17-keV neutrinos^{63,64} really do exist (in addition to CDM) and decay on a timescale of $\sim 1\text{--}5$ years⁶⁵. Alternatively, one can construct models of inflation that are tuned (for example, by involving multiple scalar fields) to produce seed fluctuations with a characteristic scale similar to that of galaxy clusters⁶⁶. A third possibility is that primordial fluctuations do not arise from quantum processes during inflation, but rather reflect the gravitational effects of topological defects generated at a different phase transition. Recent work indicates that the cosmic string model, for a long time the most promising of its kind, actually produces less large-scale structure than standard CDM⁶⁷. Cosmic textures,

a type of topological knot that unties itself as it shrinks, may provide a more successful model⁶⁸. Notice that each of these proposals requires both CDM and a flat universe; the only differences from the standard theory concern the nature of the seed fluctuations. Although some of these models can undoubtedly produce extra large-scale structure, it is not yet clear that they inherit the impressive successes of the standard model on galactic and intermediate scales. Furthermore, each of these models requires additional new physics and, except possibly for the last, parameters that are far from natural.

There may be less exotic ways of producing extra large-scale structure in the galaxy distribution which require nothing more than low-energy astrophysics. We have already noted that the relation between the distributions of galaxies and mass is uncertain in the CDM picture and is poorly constrained by current observations. Although it is a convenient model, the 'high-peak' biasing scheme may be a poor representation of where galaxies actually form. Possibly the galaxy formation process could itself introduce structure in the galaxy distribution on large scales. For example, the efficiency of star formation in a protogalaxy may depend on the binding energy of its dark-matter halo, on the physical state and magnetic field content of its gas, and on the radiation field in which it finds itself. Ultraviolet light from a quasar or massive protogalaxy can dissociate molecules and ionize protogalactic gas out to distances of tens of megaparsecs. If this (or any other large-scale coherent effect) has a significant influence on star formation efficiency, the abundance of the luminous galaxies that dominate all existing surveys of structure could be modulated on a very large scale^{69,70}. In this case, some component of the observed large-scale structure might be entirely unrelated to structure in the mass distribution.

Such uncertainties in galaxy formation limit the predictive power of the CDM theory. Nevertheless, precise predictions concerning the distribution of the dark matter can be made, and can be tested in several ways. As we have already remarked, it should be possible to use galaxy motions to probe the mass distribution on large scales. In addition, measurements of fluctuations in the brightness of the microwave background radiation should provide firm constraints on irregularities at early times.

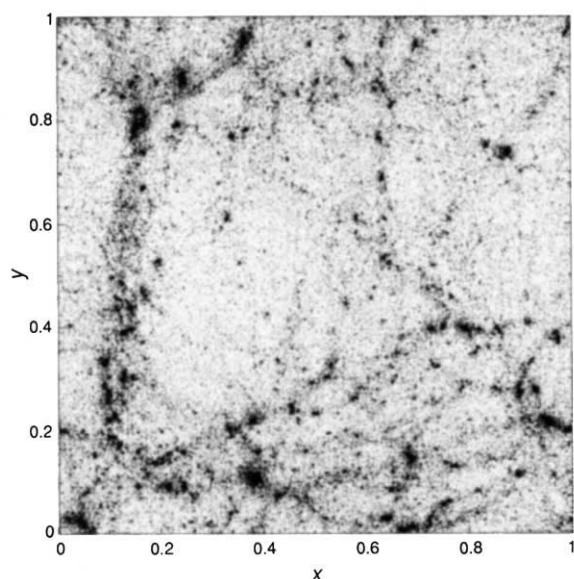


FIG. 3 Projection from a three-dimensional numerical simulation of a flat CDM universe showing the mass distribution in a cube of side 5 Mpc at a redshift $z \approx 3$. The masses of typical bound groups of particles range between a few times 10^8 solar masses ($M_\odot$) and $10^{10} M_\odot$. The small-scale structure predicted in the CDM theory leads to numerous observational tests of the model. For example, gas in the smaller systems at these redshifts could explain the Ly α absorption lines observed in quasar spectra⁷⁵ and star formation in the larger systems might render them visible at optical wavelengths.

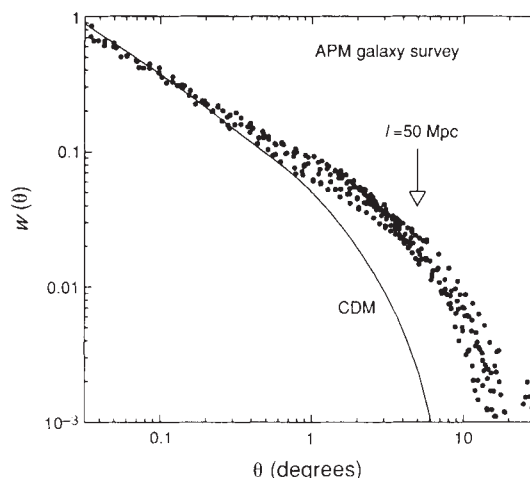


FIG. 4 Angular correlation function $w(\theta)$ for galaxies in the Automatic Plate Measuring Machine galaxy survey⁴³. The angular correlation function measures the joint probability δP of finding galaxies in each of two elements of solid angle $\delta\Omega_1$ and $\delta\Omega_2$ separated by an angle θ : $\delta P = \mathcal{N}^2(1 + w(\theta))\delta\Omega_1\delta\Omega_2$, where $\mathcal{N}$ is the mean surface density of galaxies. The dots show results for galaxies in the magnitude range 17–20.5 (ref. 43) normalized to an effective depth at which a spatial scale of 50 Mpc (for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$) subtends an angle of $\sim 5^\circ$. The curve shows the predictions of the standard CDM model if we assume that the simple 'high-peak' biasing scheme relates the galaxy distribution to fluctuations in the mass.

The COBE satellite has set impressive upper limits on such fluctuations: the background is constant to better than one part in 20,000 on all scales larger than 7° (ref. 71), but this is still well above the predictions of the standard CDM model. The best prospects for measuring the fluctuations predicted by CDM come from ground-based experiments designed to measure angular scales of $\sim 1^\circ$. These already give upper limits that are not far above the values expected in a CDM universe^{72,73} and should reach the required sensitivity within the next few years. Finally, it has recently been suggested that gravitationally induced distortions in the images of distant galaxies can be used to probe the large-scale distribution of mass, an idea which may soon be tested in practice⁷⁴.

According to the CDM model⁷⁵, galaxy formation should have peaked at redshift $z \sim 1-3$, corresponding to $\sim 10^{10}$ years ago when the Universe was $\sim 30\%$ of its present size. This might seem impressively far in the past, but until recently most cosmologists would have regarded it as unacceptably late⁷⁶. There is now a growing body of evidence^{77,78} for strong galaxy evolution suggesting that many galaxies formed recently. A recent epoch of galaxy formation is one of the clearest and most interesting predictions of the CDM theory, and we would count it as a major success for the model if observers succeeded in identifying forming galaxies at $z \approx 2$. On the other hand, a high abundance of massive galaxies at $z \geq 5-10$ would probably force us to remove the question mark from our title⁷⁹.

In a CDM universe, galaxy clustering also evolves very rapidly at recent epochs³⁸. The observational situation here is controversial. Rapid evolution in the colours and in other properties of individual galaxies in clusters is now well documented^{80,81}. Recent studies also suggest rapid evolution of the global properties of clusters selected by their X-ray emission^{82,83}. Within a few years the Rosat X-ray satellite should provide a large enough sample of such clusters to give a clear-cut test of the theoretical predictions^{38,84}. But the rapid evolution of structure at recent times has also been used as an argument against the CDM model on the grounds that galaxy collisions and mergers may occur at a rate that is incompatible with the existence of the thin stellar disks seen in most galaxies⁸⁵. More detailed work on merger rates, on the merging process and on disk formation is needed to evaluate this argument.

Many details of the CDM model, particularly those associated with galaxy formation, remain to be worked out, but the prospects seem promising. The next generation of sky surveys at optical, infrared and X-ray wavelengths, in combination with the measurement of redshifts for hundreds of thousands of galaxies, will produce a qualitative improvement in our knowledge of the large-scale galaxy distribution. If CDM is right, the 8-10-m apertures of the new generation of optical telescopes will enable us to see forming galaxies, to measure the evolution of galaxy clustering and to use absorption lines in the spectra of background quasars to study the gaseous component of protogalaxies. These observations will allow us to quantify many aspects of large-scale structure and to explore the evolution of the Universe over the latest 90% of its history. Together with the rapidly improving technology for searches of variations in the brightness of the microwave background, they provide our best chance of testing the essential elements of the CDM picture for structure formation.

Finally, CDM could be detected in laboratory experiments. Although it is unimportant for cosmogony whether CDM be gravitinos, photinos, axions or any other of the suggested candidate particles, its identity can be critical for attempts at direct detection. Investigations at large accelerators already place constraints on some dark matter candidates⁸⁶. Future experiments at the Large Electron-Positron collider, Large Hadron Collider and the Superconducting Supercollider may confirm the existence of supersymmetric particles; microwave resonance experiments are searching for axions of the required mass⁸⁷; and experiments in deep mines^{88,89} could detect dark matter

from the halo of our own Galaxy if it is any of a large number of the candidates. The attempt to identify the substance that apparently constitutes more than 90% of the mass of the Universe is surely one of the greatest challenges in contemporary physics. There are good reasons to hope that the mystery will be solved during this decade. $\square$

M. Davis is at the Astronomy and Physics Departments, University of California, Berkeley, California 94720, USA. G. Efstathiou is at the Department of Physics, University of Oxford, Mullard and Astrophysics Laboratory, Keble Road, Oxford OX1 3RH, UK. C. S. Frenk is at the Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK. S. D. M. White is at the Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK.

- Saunders, W. et al. *Nature* **349**, 32-38 (1991).
- Carr, B. J. *Comments Astrophys.* **14**, 257-280 (1990).
- Turner, M. S. *Phys. Scripta* **T36**, 167-182 (1991).
- Harrison, E. R. *Phys. Rev. D* **1**, 2726-2730 (1970).
- Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 84-89 (1970).
- Guth, A. H. *Phys. Rev. D* **23**, 347-356 (1981).
- Linde, A. D. *Phys. Lett.* **B108**, 389-393 (1982).
- Albrecht, A. & Steinhardt, P. *Phys. Rev. Lett.* **48**, 1220-1223 (1982).
- Guth, A. & Pi, S. Y. *Phys. Rev. Lett.* **49**, 1110-1113 (1982).
- Hawking, S. *Phys. Lett.* **B115**, 295-297 (1982).
- Starobinski, A. *Phys. Lett.* **B117**, 175-178 (1982).
- Bardeen, J. M., Steinhardt, P. J. & Turner, M. S. *Phys. Rev. D* **28**, 679-693 (1983).
- Colless, M. M. & Hewett, P. *Mon. Not. R. astr. Soc.* **224**, 453-472 (1987).
- Zabludoff, A. I., Huchra, J. R. & Geller, M. J. *Astrophys. J. Suppl.* **74**, 1-36 (1990).
- Olive, K. A., Schramm, D. N., Steigman, G. & Walker, T. P. *Phys. Lett.* **B236**, 454-460 (1990).
- Silk, J. *Astrophys. J.* **151**, 459-472 (1968).
- Zel'dovich, Ya. B. *Mon. Not. R. astr. Soc.* **160**, 1P-4P (1972).
- Wilson, M. L. & Silk, J. *Astrophys. J.* **243**, 14-25 (1981).
- Lyubimov, V. A., Novikov, E. G., Nozik, V. Z., Tretyakov, E. F. & Kozik, V. F. *Phys. Lett.* **B94**, 266-268 (1980).
- White, S. D. M., Frenk, C. S. & Davis, M. *Astrophys. J.* **274**, L1-L5 (1983).
- Peebles, P. J. E. *Astrophys. J.* **263**, L1-L6 (1982).
- Peebles, P. J. E. *The Large-Scale Structure of the Universe* (Princeton University Press, 1980).
- Mather, J. et al. *Astrophys. J.* **354**, L37-L40 (1990).
- Ostriker, J. P. & Cowie, L. *Astrophys. J.* **243**, L127-L132 (1981).
- Ostriker, J. P., Thompson, C. & Witten, E. *Phys. Lett.* **B180**, 231-239 (1986).
- Yahil, A. in *Large-scale Motions in the Universe* (eds Rubin, V. C. & Coyne, G. V.) 219-253 (Princeton University Press, 1988).
- Blumenthal, G. R., Faber, S. M., Primack, J. R. & Rees, M. J. *Nature* **311**, 517-525 (1984).
- Frenk, C. S. *Phys. Scripta* **T36**, 70-87 (1991).
- White, S. D. M. & Rees, M. J. *Mon. Not. R. astr. Soc.* **183**, 341-358 (1978).
- Katz, N. & Gunn, J. E. *Astrophys. J.* **377**, 365-381 (1991).
- Evrard, A. E. *Astrophys. J.* **363**, 349-366 (1990).
- Davis, M., Efstathiou, G., Frenk, C. S. & White, S. D. M. *Astrophys. J.* **292**, 371-394 (1985).
- Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. S. *Astrophys. J.* **304**, 15-61 (1986).
- White, S. D. M., Davis, M., Efstathiou, G. & Frenk, C. S. *Nature* **330**, 451-454 (1987).
- Carlberg, R. G., Couchman, H. M. P. & Thomas, P. *Astrophys. J.* **352**, L29 (1990).
- Bertschinger, E. & Gelb, J. M. *Computers Phys.* **5**, 164-179 (1991).
- Gott, J. R., Melott, A. L. & Dickinson, M. *Astrophys. J.* **306**, 341-357 (1986).
- Frenk, C. S., White, S. D. M., Efstathiou, G. & Davis, M. *Astrophys. J.* **351**, 10-21 (1990).
- Kirshner, R. P., Oemler, A., Schechter, P. L. & Shectman, S. A. *Astrophys. J.* **248**, L57-L60 (1981).
- White, S. D. M., Frenk, C. S., Davis, M. & Efstathiou, G. *Astrophys. J.* **313**, 505-516 (1987).
- Geller, M. & Huchra, J. *Science* **246**, 897-903 (1989).
- Park, C. *Mon. Not. R. astr. Soc.* **242**, 59P-61P (1990).
- Maddox, S. J., Efstathiou, G., Sutherland, W. & Loveday, J. *Mon. Not. R. astr. Soc.* **242**, 43P-47P (1990).
- Efstathiou, G. et al. *Mon. Not. R. astr. Soc.* **247**, 10P-14P (1990).
- Bahcall, N. & Soneira, R. *Astrophys. J.* **270**, 20-38 (1983).
- Peacock, J. A. & Nicholson, D. *Mon. Not. R. astr. Soc.* **253**, 307-315 (1991).
- Efstathiou, G., Dalton, G. B., Sutherland, W. & Maddox, S. J. *Mon. Not. R. astr. Soc.* (in the press).
- Dalton, J., Efstathiou, G., Sutherland, W. & Maddox, S. J. *Astrophys. J.* (in the press).
- Dressler, A. *Nature* **350**, 391-397 (1991).
- Dekel, A., Bertschinger, E. & Faber, S. M. *Astrophys. J.* **364**, 349-369 (1990).
- Bertschinger, E., Dekel, A. & Faber, S. M., Dressler, A. & Burstein, D. *Astrophys. J.* **364**, 370-395 (1990).
- Groth, J. E., Juszkiewicz, R. & Ostriker, J. P. *Astrophys. J.* **346**, 558-565 (1989).
- Kaiser, N. & Lahav, O. *Mon. Not. R. astr. Soc.* **237**, 129-162 (1989).
- Gorski, K., Davis, M., Strauss, M. A., White, S. D. M. & Yahil, A. *Astrophys. J.* **344**, 1-19 (1989).
- Rowan-Robinson, M. in *Observational Tests of Cosmological Inflation* (eds Shanks, T., Banday, A. J., Ellis, R. S., Frenk, C. S. & Wolfendale, A. W.) 161-171 (Kluwer, Dordrecht, 1991).
- Bond, J. R. & Efstathiou, G. *Astrophys. J.* **285**, L45-L48 (1984).
- Vittorio, N. & Silk, J. *Astrophys. J.* **285**, L39-L44 (1984).
- Peebles, P. J. E. *Astrophys. J.* **284**, 439-444 (1984).
- Efstathiou, G., Sutherland, W. & Maddox, S. J. *Nature* **348**, 705-707 (1990).
- Jacoby, G. H., Ciardullo, R. & Ford, H. C. *Astrophys. J.* **356**, 332-349 (1990).
- Tonry, J. L. *Astrophys. J.* **373**, L1-L4 (1991).
- Weinberg, S. *Rev. mod. Phys.* **61**, 1-23 (1989).
- Simpson, J. J. *Phys. Rev. Lett.* **54**, 1891-1893 (1985).
- Hime, A. & Jelley, N. A. *Phys. Lett.* **B257**, 441-450 (1991).
- Bond, J. R. & Efstathiou, G. *Phys. Lett.* **B265**, 245-250 (1991).
- Salopek, D. S., Bond, J. R. & Bardeen, J. M. *Phys. Rev. D* **40**, 1753-1788 (1990).
- Albrecht, A. & Stebbins, A. *Fermilab Preprint No. 91/17A* (1991).
- Turok, N. *Phys. Scripta* **T36**, 135-141 (1991).
- Babul, A. & White, S. D. M. *Mon. Not. R. astr. Soc.* **253**, 31P-34P (1991).
- Bower, R. G., Coles, P., Frenk, C. S. & White, S. D. M. *Mon. Not. R. astr. Soc.* (submitted).
- Smoot, G. et al. *Astrophys. J.* **371**, L1-L6 (1991).
- Bond, J. R., Efstathiou, G., Lubin, P. M. & Meinhold, P. R. *Phys. Rev. Lett.* **66**, 2179-2182 (1990).
- Vittorio, N., Meinhold, P., Muciaccia, P. F., Lubin, P. & Silk, J. *Astrophys. J.* **372**, L1-L4 (1991).
- Miranda-Escude, J. *Astrophys. J.* **380**, 1-8 (1991).

75. White, S. D. M. & Frenk, C. S. *Astrophys. J.* **379**, 52–79 (1991).
76. Peebles, P. J. E. in *The Epoch of Galaxy Formation* (eds Frenk, C. S. et al.) 1–13 (Kluwer, Dordrecht, 1989).
77. *The Epoch of Galaxy Formation* (eds Frenk, C. S. et al.) (Kluwer, Dordrecht, 1989).
78. Cowie, L. L. *Phys. Scripta* **T36**, 102–107 (1991).
79. Efsthathiou, G. & Rees, M. J. *Mon. Not. R. astr. Soc.* **230**, 5P–12P (1988).
80. Butcher, H. & Oemler, A. *Astrophys. J.* **219**, 18–30 (1978).
81. Dressler, A. & Gunn, J. E. in *Evolution of the Universe of Galaxies* (ed. Kron, R. G.) *Astr. Soc. Pacific Conf. Series* **10**, 200–208 (1990).
82. Edge, A. C., Stewart, G. C., Fabian, A. C. & Arnaud, K. A. *Mon. Not. R. astr. Soc.* **245**, 559–569 (1990).
83. Gioia, I. M. et al. *Astrophys. J. Suppl.* **72**, 567–619 (1990).
84. Evrard, A. E. & Henry, J. P. *Astrophys. J.* **383**, 95–103 (1991).
85. Ostriker, J. P. in *Evolution of the Universe of Galaxies* (ed. Kron, R. G.) *Astr. Soc. Pacific Conf. Series* **10**, 25–29 (1990).
86. Ellis, J. *Phys. Scripta* **T36**, 145–152 (1991).
87. Hagmann, C., Sikivie, P., Sullivan, N. S. & Tanner, D. B. *Phys. Rev. D* **42**, 1297–1300 (1990).
88. Smith, P. F. & Lewin, J. B. *Phys. Rep.* **187**, 203–280 (1990).
89. Primack, J. R., Seckel, D. & Sadoulet, B. *Ann. Rev. Nucl. Sci.* **38**, 751–778 (1988).

ACKNOWLEDGEMENTS. M.D. thanks the SERC for support during a visit to Oxford.

ARTICLES

***Caenorhabditis elegans* gene *ced-9* protects cells from programmed cell death**

Michael O. Hengartner, Ronald E. Ellis* & H. Robert Horvitz

Howard Hughes Medical Institute, Department of Biology, Room 56-629, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

The gene *ced-9* of the nematode *Caenorhabditis elegans* acts to protect cells from programmed cell death. A mutation that abnormally activates *ced-9* prevents the cell deaths that occur during normal *C. elegans* development. Conversely, mutations that inactivate *ced-9* cause cells that normally live to undergo programmed cell death; these mutations result in embryonic lethality, indicating that *ced-9* function is essential for development. The *ced-9* gene functions by negatively regulating the activities of other genes that are required for the process of programmed cell death.

ced-9 kill animals, apparently by causing cells that normally survive to activate *ced-3* and *ced-4* and die. These results indicate that *ced-9* normally acts to prevent cells that should live from undergoing programmed cell death.

A *ced-9(gf)* allele prevents cell deaths

While screening for new *C. elegans* mutations that affect programmed cell death¹⁵, we isolated and genetically characterized a dominant mutation, *n1950*, that prevents programmed cell deaths (data not shown). We mapped *n1950* to the right arm of the third chromosome, close to and about 0.05 map units to the right of the mutation *unc-69(e587)*. The *n1950* mutation defines a new gene, *ced-9 III*.

To quantify the effects of the *ced-9(n1950)* mutation on programmed cell deaths, we counted cells in the anterior half of the pharynx of *ced-9(n1950)* animals. In wild-type animals there are 49 cell nuclei in this region^{9,16}, and in *ced-3* and *ced-4* animals there are 12–14 additional nuclei (Table 1a). Similarly, in *ced-9(n1950)* animals there are about 13 extra nuclei in the anterior pharynx. These extra nuclei correspond exactly in position as well as in number to those that fail to die in *ced-3* and *ced-4* mutants.

Many extra cells survive not only in *ced-9(n1950)* homozygotes but also in *ced-9(n1950)/+* heterozygotes, indicating that the *n1950* phenotype is dominant (Table 1b). In addition, the *ced-9(n1950)* mutation has a maternal effect: about twice as many cells fail to die in heterozygotes generated by mothers carrying at least one copy of the *ced-9(n1950)* mutation than in heterozygotes generated by homozygous wild-type mothers (Table 1b), suggesting that maternal *ced-9* gene product is contributed to the developing oocyte.

Two observations indicate that *ced-9(n1950)* is a gain-of-function (*gf*) mutation. First, *n1950* is a rare mutation with dominant effects (we recovered only one allele in a screen of over 24,000 haploid genomes¹⁵, which is a frequency about 10-fold lower than that at which typical loss-of-function mutations are recovered^{17–19}). Second, a deletion (*nDf40*) that removes the *ced-9* gene does not have a dominant effect on cell death (Table 1b).

To study the effects of *ced-9(n1950)* on programmed cell deaths in regions other than the anterior pharynx, we investigated whether *n1950* could prevent the accumulation of cell corpses in *ced-1* and *ced-5* mutants. In wild-type animals, dying cells are rapidly engulfed and degraded by a neighbouring cell. In *ced-1* and *ced-5* mutants, this engulfment process is blocked,

NATURALLY occurring or programmed cell deaths have an important role in animal development and homeostasis. Such deaths, which remove cells that are not needed or that have already served their purposes, are observed in a wide variety of tissues both in vertebrates and in invertebrates^{1–3}. Many programmed cell deaths seem to require active participation by the dying cells, because inhibitors of RNA or protein synthesis block the process^{4–7}. During *C. elegans* development, 131 cells undergo programmed cell death^{8,9}; these dying cells undergo a series of morphological changes that can be recognized both by electron and light microscopy^{8,10}. Many mutations affecting various steps in this process have been identified; these mutations define a genetic pathway for programmed cell death in *C. elegans*^{11–13}. The activities of two genes, *ced-3* and *ced-4* (cell death abnormal), are necessary for programmed cell deaths to occur: mutations that inactivate either *ced-3* or *ced-4* result in the survival of almost all cells that normally die during development¹². These two genes probably function in the cells that die and might encode proteins that are themselves toxic or that regulate toxic activities¹⁴.

How are potentially lethal activities like those of *ced-3* and *ced-4* controlled, so that only the appropriate cells are killed? We have discovered a new gene, *ced-9*, that seems to regulate the activities of *ced-3* and *ced-4*. A *ced-9* gain-of-function mutation prevents cells from dying, whereas mutations that inactivate

* Present address: Laboratory of Molecular Biology, 1525 Linden Drive, University of Wisconsin at Madison, Madison, Wisconsin 53706, USA.

TABLE 1 The gain-of-function allele *ced-9(n1950)* prevents programmed cell deaths

(a) <i>ced-9(n1950)</i> prevents programmed cell deaths									
Maternal genotype		Zygotic genotype		Extra cells in anterior pharynx	No of animals				
<i>ced-3/ced-3</i>		<i>ced-3/ced-3</i>		12.5±0.7	30				
<i>ced-4/ced-4</i>		<i>ced-4/ced-4</i>		13.9±0.5	40				
<i>ced-9(n1950)/ced-9(n1950)</i>		<i>ced-9(n1950)/ced-9(n1950)</i>		13.3±0.6	45				
(b) <i>ced-9(n1950)</i> is a dominant gain-of-function mutation and shows a maternal effect									
+/+		+/+		0.03±0.05	60				
		<i>Df</i> /+		0.00	50				
		<i>ced-9(n1950)</i> /+		5.3±0.8	25				
<i>ced-9(n1950)</i> /+		+/+		0.2±0.2	25				
		<i>ced-9(n1950)</i> /+		11.4±0.8	30				
		<i>ced-9(n1950)/ced-9(n1950)</i>		13.7±0.5	30				
<i>ced-9(n1950)/ced-9(n1950)</i>		<i>ced-9(n1950)</i> /+		11.8±0.6	30				
		<i>ced-9(n1950)/ced-9(n1950)</i>		13.3±0.6	45				
(c) <i>ced-9(n1950)</i> suppresses the accumulation of cell corpses									
Genotype	Corpses pharynx		Corpses head		Corpses			Extra cells	
		<i>n</i>		<i>n</i>	P9-P11	P12	Tail	P9-P11	<i>n</i>
Wild type (N2)	0	50	0.0±0.1	50	0	0	0	0	30
<i>ced-1</i>	0.8±0.2	100*	28	10†	3.5±0.3	1.7±0.3	1.7±0.3	0.4±0.3	30
<i>ced-1; ced-3</i>	0.02±0.04	50	0.3±0.1	50	0.03±0.07	0	0.3±0.2	3.9±0.1	30
<i>ced-1; ced-4</i>	0.02±0.04	50	0.7±0.2	50	0.03±0.07	0	0.3±0.2	4.0±0.1	30
<i>ced-1; ced-9(n1950)</i>	0	30	0.5±0.3	30	0	0	0.3±0.2	4.0±0.1	30
<i>ced-5</i>	3.6±0.6	25	16±5	10	3.0±0.5	2.2±0.3	4.6±0.8	0.2±0.2	21
<i>ced-5 ced-3</i>	0.1±0.1	40	0.5±0.2	40	0.05±0.10	0	0.4±0.4	3.9±0.1	20
<i>ced-4; ced-5</i>	0.2±0.2	40	1.0±0.3	40	0.05±0.10	0	1.6±0.5	3.9±0.1	20
<i>ced-9(n1950); ced-5</i>	0.1±0.1	100	0.8±0.4	25	0	0	1.1±0.3	3.8±0.1	30
(d) <i>ced-9(n1950)</i> prevents the deaths of the HSN neurons in <i>egl-1</i> mutants									
Genotype	HSNs missing (%)		No of sides		Egg-laying defective (%)		<i>n</i>		
Wild type (N2)	1		250		0.4		704		
<i>egl-1</i>	99		200		99		447		
<i>ced-3; egl-1</i>	0		160		0.2		599		
<i>ced-4; egl-1</i>	0		100		0		417		
<i>ced-9(n1950); egl-1</i>	0		200		0		417		

We quantified cell survival by counting the cells in the procorpus and metacarpus, which together constitute the anterior half of the pharynx¹⁶. In wild-type animals there are 49 cell nuclei in this region. Cells that die are generated in characteristic positions⁹, making it easy to identify and count cells that have failed to die. *a*, The genotypes of animals studied were as shown. *b*, The complete genotypes were, from top to bottom: wild-type (N2), non-Unc progeny of *et1 unc-36/nDf40 dpy-18* males crossed with *unc-36* hermaphrodites, non-Unc progeny of *n1950* males crossed with *unc-69* hermaphrodites, Unc Dpy progeny from *+n1950/+unc-69 dpy-18* hermaphrodites, non-Lon non-Dpy progeny from *dpy-17 lon-1 +/+ +n1950 dpy-18* hermaphrodites, Unc-49 progeny from *unc-69 +/+ +n1950 unc-49* heterozygous hermaphrodites, non-Unc progeny of wild-type (N2) males crossed with *unc-69 n1950* hermaphrodites, and *n1950* self-progeny from *n1950* homozygous hermaphrodites. *c*, For the pharyngeal and head corpses, we scored only young larvae with four cells in the gonad, that is, between hatching and the middle of the first larval stage²⁴. For ventral cord corpses (descendants from the blast cells P9-P12) and for tail corpses, we scored third larval stage animals. Extra cells, number of extra cells among the descendants of P9, P10, and P11. The divisions of these blast cells generate four programmed cell deaths in the wild type (see text and Fig. 2a). *n*, Number of animals scored; * Data from ref. 13. † Data from ref. 12. *d*, HSN missing (%), per cent of missing or grossly displaced HSN neurons. Only first or second larval stage animals were scored. No of sides, number of sides scored (there are two HSNs per animal, one on each side²¹). To score egg laying, staged worms were grown at 20 °C. Animals were observed using a dissecting microscope on the second day of adulthood, and those bloated with late-stage eggs were considered egg-laying-defective²⁰. *n*, Number of animals scored. The alleles used were: *ced-1(e1735)*, *ced-3(n1717)*, *ced-4(n1162)*, *ced-5(n1812)*, *ced-9(n1950)*, *dpy-17(e164)*, *dpy-18(e364)*, *egl-1(n487sdts)*, *lon-1(e1820)*, *unc-36(e251)*, *unc-49(e382)*, and *unc-69(e587)*. *et1(e873)*, a translocation chromosome with a breakpoint that disrupts *unc-36* gene function, prevents crossing over on the right arm of chromosome III (ref. 37). *nDf40* is a new deficiency we isolated as a *cis*-acting suppressor of *n1950* (see text). Methods. Animals were anaesthetized with 30 mM NaN₃ (ref. 38) and observed using Nomarski optics microscopy⁸. Average numbers are shown with, if appropriate, their 95% confidence limits, determined by the *t*-test using the *StatView II* program (Abacus Concepts, Berkeley, California).

leading to an accumulation of undegraded cell corpses that can easily be seen in young larvae^{11,13}. Mutations that inactivate *ced-3* or *ced-4* block programmed cell death and therefore prevent the accumulation of dead cells in *ced-1* or *ced-5* animals (ref. 12 and our unpublished observations). Similarly, very few corpses appear anywhere in *ced-1; ced-9(n1950)* or *ced-9(n1950); ced-5* double mutants (Table 1c). Thus, the *ced-9(n1950)* mutation, like mutations in *ced-3* and *ced-4*, prevents programmed cell deaths throughout the animal.

We also studied the effects of *ced-9(n1950)* on the survival and function of a specific pair of nerve cells, the HSNs (hermaphrodite-specific neurons)²⁰⁻²³. The two HSN neurons innervate the vulval muscles and control egg-laying by hermaphrodites. Mutations in the gene *egl-1* cause these cells to undergo programmed cell death, resulting in egg-laying defective animals^{12,20,23}. Mutations in *ced-3* and *ced-4*, which block pro-

grammed cell death, prevent the HSNs from dying in *egl-1* mutants and suppress the egg-laying defect¹². Similarly, the HSNs are present in *ced-9(n1950); egl-1* double mutants, and egg-laying by these animals is normal (Table 1d). Thus, *ced-9(n1950)*, like *ced-3* and *ced-4* mutations, not only prevents cells from dying but, at least in this case, also generates surviving cells that are sufficiently healthy to function.

Isolation of *ced-9(lf)* mutations

Because the *ced-9(n1950)* mutation causes a gain of gene function (see above), we sought to isolate mutations that reduce or eliminate *ced-9* activity (*ced-9* loss-of-function (*lf*) mutations) by screening for *cis*-dominant suppressors of *ced-9(n1950)*: we expected that a second mutation in *ced-9* could suppress the dominant effects of *n1950* by inactivating *ced-9* (Fig. 1). After screening 9,000 haploid genomes, we isolated three candidate

TABLE 2 Phenotypes of *ced-9(lf)* mutants

	<i>ced-9(+)</i>	<i>ced-9(+)</i>	<i>n1950 n2161</i>			<i>n1950 n2161</i>	<i>n1950 n2161</i>	<i>n1950 n2161</i>	<i>n1653ts</i>	<i>n1653ts</i>	<i>n1653ts</i>	<i>n1950 n2077</i>	<i>n1950 n2077</i>
Genotype*	20 °C	Df 20 °C	15 °C	20 °C	23 °C	25 °C	Df 20 °C	<i>n1950 n2077</i> 20 °C	25 °C	Df 25 °C	<i>n1950 n2077</i> 25 °C	20 °C	Df 20 °C
(a) Sterility and maternal-effect lethality													
Eggs laid per animal	209 ± 33	202 ± 60	117 ± 36	97 ± 31	45 ± 22	6.3 ± 4.5	23 ± 14	40 ± 9	2.7 ± 1.1	0	0.3 ± 0.4	1.6 ± 1.4	0.8 ± 1.6
Hatching (%)	99 ± 1	75 ± 2	12 ± 3	2.4 ± 0.7	0.4 ± 0.4	0	0	0	38 ± 24	NA	0.7 ± 1.4	0	0
L1 arrest (%)	0	13 ± 3	100	100	100	NA	NA	NA	40 ± 36	NA	100	NA	NA
	<i>n</i> = 14	<i>n</i> = 9	<i>n</i> = 23	<i>n</i> = 42	<i>n</i> = 36	<i>n</i> = 60	<i>n</i> = 50	<i>n</i> = 49	<i>n</i> = 26	<i>n</i> = 15	<i>n</i> = 30	<i>n</i> = 20	<i>n</i> = 24
(b) Egg-laying defect													
Egg-laying defective (%)	0	0	64 ± 16	76 ± 13	94 ± 7	98 ± 3	96 ± 4	96 ± 6	NA†	NA†	NA†	NA†	NA†
HSNs missing (%)	0	0	77	87	94	95	95	ND	ND	ND	ND	99	100
	<i>n</i> = 100	<i>n</i> = 48	<i>n</i> = 118	<i>n</i> = 138	<i>n</i> = 100	<i>n</i> = 130	<i>n</i> = 60	<i>n</i> = 40				<i>n</i> = 220	<i>n</i> = 42
(c) Absence of rays in male tails													
Rays per side	8.9 ± 0.1	8.6 ± 0.2	8.0 ± 0.3	6.6 ± 0.3	5.9 ± 0.3	5.4 ± 0.4	6.0 ± 0.3	5.9 ± 0.3	8.6 ± 0.2	7.6 ± 0.3	8.1 ± 0.3	4.6 ± 0.3	4.9 ± 0.6
	<i>n</i> = 68	<i>n</i> = 34	<i>n</i> = 40	<i>n</i> = 40	<i>n</i> = 58	<i>n</i> = 34	<i>n</i> = 62	<i>n</i> = 38	<i>n</i> = 38	<i>n</i> = 44	<i>n</i> = 52	<i>n</i> = 81	<i>n</i> = 26

a, Numbers of eggs laid by first generation *ced-9(lf)* hermaphrodites and the stages of development at which the progeny of these hermaphrodites arrested. First generation hermaphrodites were transferred to fresh plates every 12 h, and the number of eggs they laid were counted. Note that the absence of HSNs retards but does not prevent egg-laying²⁰, so that the sterility observed in *ced-9(lf)* animals (as reflected by the number of eggs laid per animal) cannot be only an effect of the defect in egg release. For example, *egl-1(n487sdts)* animals, which are egg laying-defective because they lack HSNs, nonetheless lay an average of 204 eggs²³. Egg-laying defective *ced-9(lf)* animals do, however, fertilize a few eggs that are never laid. For example, although wild-type hermaphrodites lay all fertilized eggs within 4 days of reaching adulthood, by the seventh day *ced-9(n1950 n2077)* hermaphrodites still had 1.7 ± 1.3 eggs (number of broods = 12) remaining *in utero*, and egg-laying defective *ced-9(n1950 n2161)* hermaphrodites had 30 ± 23 eggs (number of broods = 6). The number of eggs laid by *ced-9(lf)* animals therefore usually slightly underestimates actual brood size. Hatching (%), per cent of eggs laid that hatched within 48 h of removal of the mother (wild-type eggs hatch about 14 h after fertilization⁹). L1 arrest (%), per cent of hatched progeny that failed to develop past the first (L1) larval stage within 6 days of hatching (wild-type larvae remain in the L1 stage for about 12 h⁸). *n*, Number of broods analysed. b, Per cent of animals defective in egg-laying was scored as in Table 1. Note however that for some genotypes (marked †), a significant fraction of the animals could not be scored accurately for egg-laying capability because of the small number of eggs they generated. *n*, Number of animals scored. Egg-laying defective *ced-9(lf)* animals do lay eggs in the presence of serotonin (data not shown; assayed as in ref. 20), suggesting that the serotonergic HSN neurons are defective or absent. HSN missing (%), percent of missing or grossly displaced HSN neurons, scored as in Table 1. *n*, Number of sides scored. c, Rays per side, number of rays present per side in the male tail. Young adult males were anaesthetized in 30 mM NaCl₃, placed on their backs and observed using Nomarski optics. *n*, Number of sides scored. Confidence limits (95%) were determined as in Table 1. NA, Not applicable. * All strains were homozygous (*nDf40* strains were hemizygous) for the closely linked mutation *unc-69(e587)*, which facilitates identification of the chromosome carrying the *ced-9* mutation. All *ced-9(lf)* mutations were maintained as heterozygous stocks balanced by the chromosome III balancer *qC1* (J. Austin and J. Kimble, personal communication). The *nDf40* chromosome was marked with *dpy-18(e364)*. *nDf40* fails to complement both *ced-9* and *unc-69* (see legend to Fig. 1). For the *n2161/n2077* and *n1653/n2077* trans-heterozygotes, the maternally-inherited *n2077* chromosome was marked with the *dpy-18(e364)* mutation to distinguish self- from cross-progeny. The HSN counts for the *n2161/n2077* and *n1653/n2077* genotypes were not determined because of the difficulty of scoring the Dpy phenotype in early larvae. The *ced-9(+)/Df* larvae that arrested as L1s did so as a consequence of the *unc-69(e587)* mutation, which decreases brood size and results in an incompletely penetrant L1-arrest phenotype when heterozygous with *nDf40*; by contrast, *nDf40/unc-69(+)* animals do not arrest development as L1 larvae (data not shown).

suppressor mutations tightly linked to *ced-9(n1950)*. One of these mutations, *nDf40*, behaved genetically as a large deletion (see legend to Fig. 1), indicating that our screen should have allowed the isolation of mutations that completely inactivate *ced-9*. The other two mutations, *n2077* and *n2161*, seem likely to be *ced-9* loss-of-function alleles: these two mutations failed to complement each other while complementing recessive mutations in all known genes in this region, mapped within 0.1 map units of the original *n1950* mutation (data not shown), and were obtained at a frequency of about 3×10^{-4} per haploid genome, which is comparable to that for loss-of-function mutations in other *C. elegans* genes¹⁷⁻¹⁹.

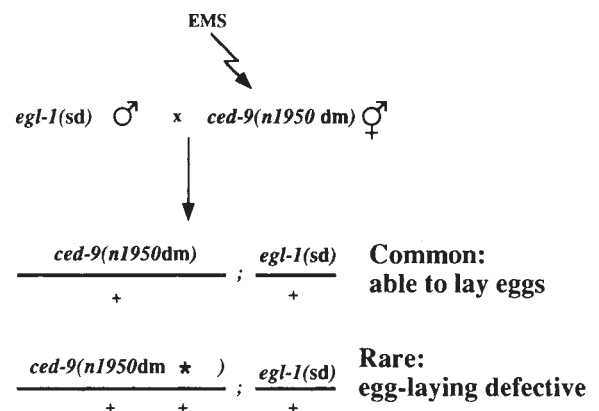
We then determined that another mutation, *n1653ts*, previously isolated in an unrelated screen for mutants with displaced or missing HSN neurons²², was also a *ced-9(lf)* allele.

We showed that *n1653* is allelic to *n2077* and *n2161* based on its position on the genetic map, the similarity of its phenotype at restrictive temperature to the phenotypes of *n2077* and *n2161* mutants, and its failure to complement *n2077* and *n2161* (data not shown). Programmed cell deaths occurred normally in *n1950 n2077/++* and *n1950 n2161/++* animals, but were blocked in *n1653/n1950* trans-heterozygotes (data not shown). The results of this *cis-trans* test demonstrate that the allelic mutations *n2077*, *n2161* and *n1653* are in the *ced-9* gene, which is defined by the mutation *n1950*, rather than in a closely linked gene.

The *ced-9(lf)* alleles cause ectopic cell deaths

Animals homozygous for *ced-9(lf)* mutations show several defects (Table 2, Fig. 2). Most obviously, homozygous *ced-9(lf)*

FIG. 1 Screen for mutations that result in a loss of *ced-9* function. The semidominant mutation *egl-1(n487sd)* causes the two HSN neurons to die by programmed cell death, so that the animal bloats with eggs^{12,20}. Because *ced-9(n1950)* dominantly suppresses *egl-1(n487)* by preventing the deaths of the HSN neurons, only animals that do not have *ced-9(n1950)* function will bloat with eggs as a result of the *egl-1* mutation. We screened for such egg-laying defective animals by mating *egl-1(n487)* V males either with *unc-69(e587) ced-9(n1950)* III; *unc-10(e102) xol-1(y9) dpy-6(e14)* X hermaphrodites or with *unc-69(e587) ced-9(n1950) dpy-18(e364)* III; *lon-2(e678) xol-1(y70)* X hermaphrodites. Egg laying-defective cross-progeny were picked and their progeny examined for any unusual phenotype. The *xol-1* mutation on the X chromosome causes male lethality³⁹ and so prevents mating among F₁ animals, which would complicate genetic analysis of new mutations. The *unc-69*, *dpy-18*, *lon-2*, *unc-10* and *dpy-6* mutations were used as closely linked genetic markers to identify the chromosomes carrying the *ced-9* and *xol-1* mutations. General genetic methods and techniques for mutagenesis with ethyl methanesulphonate are described in ref. 17. Two-factor mapping experiments showed the new mutations *n2077* and *n2161* to be tightly linked to *ced-9(n1950)* (data not shown). This screen also generated *nDf40*, a deficiency that fails to complement *unc-69*, *ced-9*, *unc-49* and several adjoining genes (data not shown). Mapping data for the *ced-9* alleles will be published elsewhere. The loss-of-function mutation *ced-9(n1950 n2077)* complements the nearby mutations *unc-*



50(e306), *ooc-4(e2078)* and *emb-25(g45ts)*; *ced-9(n1653ts)* complements *unc-69(e587)*. The *ooc-4* mutation causes a defect in oogenesis, resulting in hermaphrodite sterility (T. Doniach and J. Hodgkin, personal communication). The mutation *emb-25(g45ts)* is described in ref. 40. All other mutations are described in ref. 17.

mutants derived from *ced-9(lf)/+* heterozygous mothers hatch and grow to normal size, but generate very few eggs (partial sterility), all of which eventually die, usually during embryogenesis (maternal effect lethality). Furthermore, such first generation *ced-9(lf)* animals lack many cells normally present in wild-type animals, resulting in a number of additional defects. For example, many ventral cord motor neurons involved in the control of movement are missing, resulting in uncoordinated body movement. The HSN neurons are missing in hermaphrodites, causing an egg-laying defect (Table 2). Similarly, cells

are absent from the male tail, resulting in missing or deformed rays (Fig. 2e, f; Table 2). Furthermore, several neurons are sometimes missing from the lumbar ganglion, although their absence does not result in an obvious behavioural abnormality (data not shown).

To determine why cells are missing, we observed *ced-9(lf)* animals as they developed. The pattern of cell divisions in wild-type *C. elegans* is highly reproducible among individuals, and deviations from the normal cell lineage can be identified^{18,9,24,25}. Our studies revealed that many cells that normally

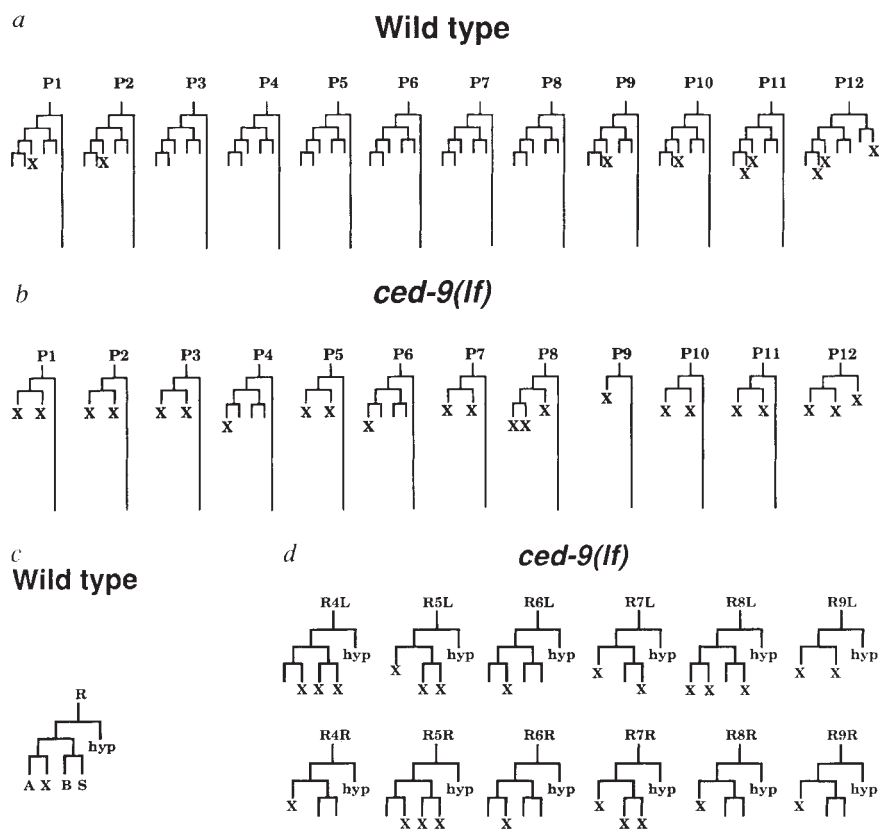


FIG. 2 The loss of *ced-9* function results in ectopic cell deaths. **a**, Cell lineages of the ventral cord blast cells P1–P12 in the wild type⁸. X, Programmed cell death. **b**, Cell lineages of P1–P12 in a *ced-9(n1950 n2077)* hermaphrodite progeny of a *qC1/unc-69 ced-9(n1950 n2077)* heterozygous mother. The exact pattern of cell deaths varied slightly among the three animals we studied. The lineages shown for P1–P10 were obtained from one animal, and those for P11 and P12 from a second. **c**, Ray lineages in the wild type^{8,25}. R, Ray precursor cell; A, type A neuron; B, type B neuron; S, ray structural cell; hyp, hypodermal cell. This pattern of division is used to form all 18 rays⁸. **d**, Cell lineages of the left and right R cells, R4L–R9L and R4R–R9R, respectively, observed in a single *ced-9(n1950 n2077)* male progeny of a *qC1/unc-69 ced-9(n1950 n2077)* heterozygous mother crossed with males of identical genotype. The left and right R1–R3 cell lineages were not followed in this particular animal. As in the ventral cord lineages, the exact pattern of ectopic cell deaths varied among the three animals studied. **e**, *unc-69* male tail. Each side has nine rays (arrowheads). **f**, *unc-69 ced-9(n1950 n2077)* male tail. This particular animal has only three rays on the left side and five on the right side (arrowheads). **g**, *ced-4 unc-69 ced-9(n1950 n2077)* male tail, showing 18 rays (arrowheads). **h**, Wild-type embryo, 350 min after fertilization, ventral view. Arrow points to a cell corpse. **i**, *ced-9(n1950 n2161)* embryo generated by a *ced-9(n1950 n2161)* mother, 350 min after fertilization, ventral view. Many corpses can be seen (arrows indicate a few). Scale bars, 10 μ m.

METHODS. Cell lineages were followed using Nomarski optics microscopy⁸. Four-dimensional-microscopy of embryos: freshly fertilized embryos were mounted for observation on 5% agar pads in a drop of M9 or egg salts⁹. Pictures of the developing embryos were taken in 18 focal planes (roughly 1 μ m apart) at 30-s intervals using an apparatus developed by J. G. White (personal communication) and stored on a 12-inch optical video disk for subsequent analysis.

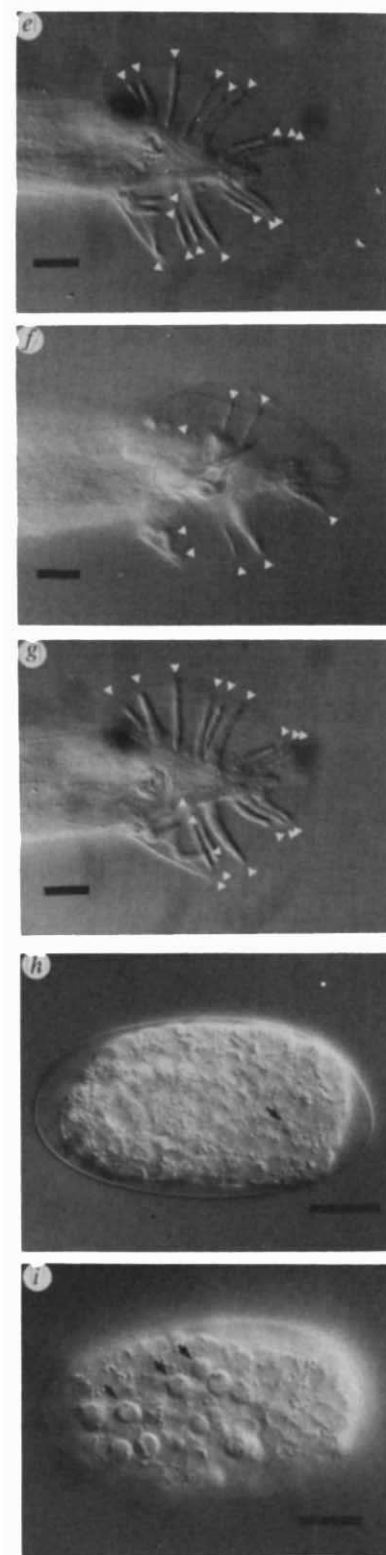


TABLE 3 Mutations in *ced-3* and *ced-4* suppress the defects resulting from the loss of *ced-9* function

Genotype*	Sterility and maternal-effect lethality			Egg-laying defect		Male tail	
	Eggs laid per animal	Viable progeny	<i>n</i>	Egg-laying defective (%)	<i>n</i>	Rays per side	<i>n</i>
<i>ced-9(+)</i>	207 ± 36	207 ± 33	14	0	100	8.9 ± 0.1	68
<i>ced-9(n1950 n2077)</i>	1.6 ± 1.4	0	20	NA†	100	4.6 ± 0.3	81
<i>ced-4 ced-9(n1950 n2077)</i>	200 ± 19	160 ± 20	12	1	84	9.0	42
<i>ced-4</i>	182 ± 17	148 ± 17	12	0	100	9.0 ± 0.1	30
<i>+ ced-9(n1950 n2077)</i>	56 ± 20	3.9 ± 2.0	7	62	71	7.4 ± 0.3	50
<i>ced-4 ced-9(n1950 n2077)</i>	154 ± 29	94 ± 22	13	0	100	9.0	28
<i>ced-9(n1950 n2077); ced-3</i>	178 ± 51	146 ± 46	6	0	66	8.9 ± 0.1	50
<i>ced-3</i>	132 ± 62	17 ± 6	10	18	40	8.7 ± 0.1	50
<i>ced-9(n1950 n2077); ced-3/+</i>							

The numbers of eggs laid were determined as described for Table 2. Viable progeny, number of progeny that grew to the fourth larval (L4) stage within 10 days of hatching (this value includes a few animals that developed from eggs that hatched internally); wild-type larvae reach the L4 stage within 2 days⁸. *n*, Number of broods analysed. See text and Table 2 for explanations of the egg-laying and ray defects. Confidence limits (95%) were determined as in Table 1. Note that *ced-3* and *ced-4* are able to suppress the *ced-9(lf)* zygotic defects in a semidominant fashion: animals homozygous for *ced-9(n1950 n2077)* but carrying only one wild-type copy of either the *ced-3* or *ced-4* genes showed milder zygotic defects than did animals with two wild-type copies of both genes, suggesting that lowering *ced-3* or *ced-4* activity can compensate for lower levels of *Ced-9* in first generation *ced-9(lf)* animals. However, one copy of *ced-3* or *ced-4* is not sufficient to suppress the maternal-effect lethality: all the viable progeny generated from *ced-9(lf)*; *ced-3/+* mothers were homozygous for the *ced-3* mutation. We also constructed double mutants between *ced-9(n1950 n2077)* and *ced-3(n1718)*, *ced-3(n1040)*, *ced-3(n1128)*, *ced-3(n1949)*, *ced-4(n1894)*, or *ced-4(n1920)*, and between *ced-9(n1950 n2161)* and *ced-3(n1717)*, *ced-4(n1162)*, *ced-4(n1894)*, or *ced-4(n1920)*. All of these double mutants were both viable and fertile, showing that the suppression of the *ced-9(lf)* defects by *ced-3* and *ced-4* is not allele-specific.

* All strains are homozygous for the closely linked mutation *unc-69(e587)* (see legend to Table 2).

† Many animals could not be scored accurately for egg-laying (see legend to Table 2).

survive in wild-type animals instead undergo programmed cell death in *ced-9(lf)* animals.

For example, we followed the divisions of the 12 ventral cord blast cells P1–P12 (collectively called Pn) (Fig. 2a, b). During the first larval (L1) stage, each P cell divides to generate an anterior daughter (Pn.a) that is a neuroblast and a posterior daughter (Pn.p) that is a hypodermal blast cell. The Pn.a cells then follow identical patterns of divisions to generate motor neurons involved in locomotion^{8,21,26}. We observed numerous ectopic cell deaths in all Pn.a lineages of *ced-9(lf)* animals, and frequently all descendants of the Pn.a neuroblasts died (Fig. 2b).

We also observed ectopic programmed cell death in the ray lineages. Rays are simple sensory structures located in the male tail, which is used for copulation. Each of the 18 rays arises from a single ray precursor cell (refs 8, 25 and Fig. 2c). Many ectopic cell deaths occurred in the ray lineages of *ced-9(lf)* males (Fig. 2d). These ectopic deaths often eliminated the ray structural cell, which is required for ray formation⁸. Thus, these deaths account for the absence of rays in *ced-9(lf)* males.

To determine the cause of the maternal-effect lethality of *ced-9(lf)* mutations, we studied the embryonic cell lineages of the progeny of *ced-9(lf)* animals. Embryos generated by mothers homozygous for the weak allele *ced-9(n1950 n2161)* usually arrested during the early stages of embryo elongation (about 450 min after fertilization⁹), although there was some variation from animal to animal. These embryos developed normally to about the 200-cell stage, at which point extensive ectopic cell deaths began to appear (Fig. 2i). These ectopic cell deaths were morphologically similar to the cell deaths that occur during normal *C. elegans* development and that also first appear at this stage⁹. Using a 'four-dimensional'-microscope (which allows time-lapse recording of multiple focal planes of a specimen; J. White, personal communication), we analysed the cell lineage of a single *ced-9(n1950 n2161)* embryo and identified 49 of the cells that died (more than 100 cells died eventually; data not shown). Of these 49 dying cells, 45 normally survive in the wild type. These 45 ectopic deaths prevented the generation of 78 cells, 50 of which would have been neurons or glial cells. Mothers of hypodermal cells and of muscle cells, as well as of a few postembryonic ectodermal blast cells, also died. We did not discern any obvious pattern to the ectopic cell deaths. Many of these deaths involved cells that in the wild type do not generate

any descendants that die. Therefore, these deaths were not simply consequences of premature activation of the pathway for programmed cell death.

We also studied embryos from mothers homozygous for the strong allele *ced-9(n1950 n2077)*. The *n2077* mutation probably results in a complete or nearly complete inactivation of *ced-9*, because *n2077* behaves like the *ced-9* deletion *nDf40* when placed in *trans* to each of the *ced-9* alleles (Table 2 and data not shown). Surprisingly, the defects and terminal phenotype associated with this allele were quite different from those of *n1950 n2161* embryos. The *F₂* *n1950 n2077* embryos arrested much earlier in development, with different individuals having from a few dozen to a few hundred cells at most. The embryos invariably looked sick, with swollen cells and abnormal granules in the cytoplasm. Furthermore cell divisions were slow and asynchronous. In those rare animals that developed sufficiently far before arresting, cell corpses started to appear at about the same stage as in *n1950 n2161* embryos. We followed the lineage of a single *ced-9(n1950 n2077)* embryo with the four-dimensional-microscope. This embryo arrested with 57 cells. Nothing resembling a programmed cell death was observed. However, blocks in mitosis and cytokinesis were apparent, with incomplete cytokineses resulting in the formation of several binucleate cells. We do not know whether these defects in cell divisions and the general sickness are effects of a lack of *ced-9* function in the embryo or are secondary consequences of abnormalities in the maternal germline. Because all of these defects are completely suppressed by mutations in *ced-3* or *ced-4* (see below), it seems likely that they are a consequence of the *ced-9(n1950 n2077)* allele rather than of another mutation carried in this strain. We suspect that these defects are caused by the ectopic activation of the pathway for programmed cell death in the maternal germline. Alternatively, the three genes *ced-9*, *ced-3* and *ced-4* might act not only in programmed cell death but also in an aspect of early *C. elegans* development that is unrelated to programmed cell death.

The *ced-9* gene antagonizes *ced-3* and *ced-4*

If the defects associated with a loss of *ced-9* function are caused entirely by the aberrant activation of the programmed cell death pathway, then mutations that prevent the process of programmed cell death might be able to suppress these defects. To test

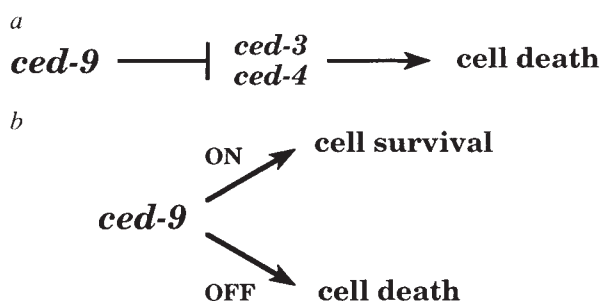


FIG. 3 Model for *ced-9* function. *a*, *ced-9* is a negative regulator of *ced-3* and *ced-4* activity. *b*, *ced-9* acts as a binary switch to regulate programmed cell death. When *ced-9* is active, the activities of *ced-3* and *ced-4* are blocked, and the cell survives. When *ced-9* is inactive, *ced-3* and *ced-4* are activated, leading to programmed cell death.

this hypothesis, we constructed double mutants using *ced-9(n1950 n2077)* and mutations in either *ced-3* or *ced-4*, two genes required for programmed cell death¹². Mutations in *ced-3* or *ced-4* completely suppressed all defects observed in *ced-9(n1950 n2077)* animals (Table 3, Fig. 2g, and unpublished observations). We obtained similar results for *ced-9(n1950 n2161)* and *ced-9(n1653ts)* (data not shown). These observations suggest that the defects seen in *ced-9(lf)* animals are indeed caused by the activation of the programmed cell death pathway. Furthermore, if these three genes are part of a regulatory pathway, these results indicate that *ced-9* acts before *ced-3* and *ced-4*, because the activities of these genes are required for *ced-9(lf)* mutations to have their effects.

Discussion

The gene *ced-9* seems to act by preventing programmed cell death, probably by antagonizing the activities of *ced-3* and *ced-4* (Fig. 3a). If *ced-9* is inactive, as in *ced-9(lf)* mutants, cells that should survive instead die by a mechanism that requires *ced-3* and *ced-4*. By contrast, in *ced-9(n1950)* mutants, cells that should die instead survive. Because the *1950* mutation has an effect opposite to that of mutations that eliminate *ced-9* function, we suspect that *n1950* results in a constitutive or over-active *ced-9* gene product. These results indicate that *ced-9* acts as a binary switch to regulate programmed cell death (Fig. 3b). Remarkably, it seems that many and possibly all cells that survive during C.

elegans development do so because *ced-9* gene activity prevents them from undergoing programmed cell death.

That both the dominant and the recessive *ced-9* mutations show maternal effects suggests that maternally-derived *ced-9* product is contributed to the egg. The amount of wild-type *ced-9* product contributed by heterozygous mothers to homozygous *ced-9(lf)* embryos seems to be sufficient to allow these embryos to survive and develop almost normally. As a consequence of this maternal rescue, the lethality that results from an absence of *ced-9* function during early development is apparent only in the second generation. Most of the ectopic cell deaths observed in the first generation of homozygous *ced-9(lf)* animals occur late, during postembryonic development. It is possible that these late lineages are more seriously affected because dilution or degradation has reduced the amount of maternal *ced-9* product to a level at which it cannot efficiently antagonize the activities of *ced-3* and *ced-4*.

How might the *ced-9* gene product (Ced-9) function to protect cells against programmed cell death? In other organisms, a variety of growth factors prevent cells from undergoing a suicidal death process (reviewed in ref. 27), and it is possible that *ced-9* controls such a growth factor or its receptor. But unlike known growth factors, Ced-9 seems to affect most or even all cells throughout development. Alternatively, *ced-9* could encode a protein that acts in the absence of any intercellular signal to block the activities of the *ced-3* and *ced-4* genes.

A number of genes, of both viral and cellular origin, can prevent programmed cell death²⁸⁻³⁰. One of these genes, the human oncogene *bcl-2*, shows intriguing similarities to *ced-9*. Overexpression of *bcl-2* prevents or delays the onset of apoptotic cell death both in B cells^{28,31} and in T cells^{32,33}. These cell deaths seem to be suicides, as gene expression is required for this process³⁴. In many tissues in which homeostasis is regulated by cell death, *bcl-2* expression occurs in progenitor and long-lived cells³⁵. These results suggest that, like *ced-9*, the *bcl-2* gene negatively regulates active cell death processes and is required in cells that survive to protect them from programmed cell death. Thus, active continuous protection from programmed cell death might be a widespread and important phenomenon.

It is possible that of the many human disorders characterized by extensive cell deaths (such as neurodegenerative diseases, stroke, and traumatic brain injury, for example see ref. 36), some are caused by processes that inactivate or bypass the functions of genes like *bcl-2* and *ced-9*. If so, an understanding of how *ced-9* acts might provide insights into the nature of these disorders. □

Received 13 January; accepted 16 March 1992.

- Glücksman, A. A. *Biol. Rev. Cambridge Phil. Soc.* **26**, 59-86 (1950).
- Saunders, J. *Science* **154**, 604-612 (1966).
- Ellis, R. E., Yuan, J. & Horvitz, H. R. *A. Rev. Cell Biol.* **7**, 663-698 (1991).
- Lockshin, R. A. *J. Insect. Physiol.* **15**, 1505-1516 (1969).
- Cohen, J. J. & Duke, R. C. *J. Immun.* **132**, 38-42 (1984).
- Martin, D. P. et al. *J. Cell Biol.* **106**, 829-844 (1988).
- Schwartz, L. M., Kosz, L. & Kay, B. K. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6594-6598 (1990).
- Sulston, J. E. & Horvitz, H. R. *Dev. Biol.* **56**, 110-156 (1977).
- Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev. Biol.* **100**, 64-119 (1983).
- Robertson, A. & Thomson, N. *J. Embryol. exp. Morph.* **67**, 89-100 (1982).
- Hedgecock, E. M., Sulston, J. E. & Thomson, J. N. *Science* **220**, 1277-1279 (1983).
- Ellis, H. M. & Horvitz, H. R. *Cell* **44**, 817-829 (1986).
- Ellis, R. E., Jacobson, D. M. & Horvitz, H. R. *Genetics* **129**, 79-94 (1991).
- Yuan, J. Y. & Horvitz, H. R. *Dev. Biol.* **138**, 33-41 (1990).
- Ellis, R. E. & Horvitz, H. R. *Development* **112**, 591-603 (1991).
- Albertson, D. G. & Thomson, J. N. *Phil. Trans. R. Soc. B275*, 299-325 (1976).
- Brenner, S. *Genetics* **77**, 71-94 (1974).
- Meneely, P. M. & Herman, R. K. *Genetics* **92**, 99-115 (1979).
- Greenwald, I. S. & Horvitz, H. R. *Genetics* **96**, 147-164 (1980).
- Trent, C., Tsung, N. & Horvitz, H. R. *Genetics* **104**, 619-647 (1983).
- White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. B311*, 1-340 (1986).
- Desai, C., Garriga, G., McIntire, S. L. & Horvitz, H. R. *Nature* **336**, 638-646 (1988).
- Desai, C. & Horvitz, H. R. *Genetics* **121**, 703-721 (1989).
- Kimble, J. & Hirsh, D. *Dev. Biol.* **70**, 396-417 (1979).
- Sulston, J. E., Albertson, D. G. & Thomson, J. N. *Dev. Biol.* **78**, 542-576 (1980).

- Chalfie, M. & White, J. in *The Nematode Caenorhabditis elegans* (eds Wood, W. B. et al.) 337-391 (Cold Spring Harbor Laboratory Press, New York, 1988).
- Barde, Y. A. *Neuron* **2**, 1525-1534 (1989).
- Vaux, D. L., Cory, S. & Adams, J. M. *Nature* **335**, 440-442 (1988).
- Henderson, S. et al. *Cell* **65**, 1107-1115 (1991).
- Clem, R. J., Fehseimer, M. & Miller, L. K. *Science* **254**, 1388-1390 (1991).
- Núñez, G. et al. *J. Immun.* **144**, 3602-3610 (1990).
- Sentman, C. L., Shutter, J. R., Hockenbery, D., Kanagawa, O. & Korsmeyer, S. J. *Cell* **67**, 879-888 (1991).
- Strasser, A., Harris, A. W. & Cory, S. *Cell* **67**, 889-899 (1991).
- Duke, R. C. & Cohen, J. J. *Lymphokine Res.* **5**, 289-299 (1986).
- Hockenbery, D. M., Zutter, M., Hickey, W., Naim, M. & Korsmeyer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6961-6965 (1991).
- Choi, D. W. *Neuron* **1**, 623-634 (1988).
- Rosenbluth, R. E. & Baillie, D. L. *Genetics* **99**, 415-428 (1981).
- Avery, L. & Horvitz, H. R. *Cell* **51**, 1071-1078 (1987).
- Miller, L. M., Plenefisch, J. D., Casson, L. P. & Meyer, B. J. *Cell* **55**, 167-183 (1988).
- Cassada, R., Isnenghi, E., Culotti, M. & von Ehrenstein, G. *Dev. Biol.* **84**, 193-205 (1981).

ACKNOWLEDGEMENTS. We thank G. Garriga, J. Hodgkin, E. Jorgensen, J. Kaplan, J. Kimble, S. C. Kim, M. Labouesse, S. Shaham, J. White, J. Yuan and members of our laboratory for discussions and comments. G. Garriga for suggesting some experiments and for the isolation of *ced-9(n1653ts)* and J. White for advice and access to his four-dimensional microscope. This work was supported by the US Public Health Service and by the Howard Hughes Medical Institute. M.O.H. was supported by a 1967 Science and Engineering Fellowship from NSERC Canada, a B1 Fellowship from FCAR Québec and a Fellowship from the Whitehead Institute for Biomedical Research. R.E.E. was supported by a National Science Foundation Graduate Fellowship and by an NIH Training Grant. H.R.H. is an Investigator of the Howard Hughes Medical Institute.

Evidence from zinc abundances for dust fractionation in chemically peculiar stars

H. Van Winckel*†, J. S. Mathis‡§ & C. Waelkens*

* Instituut voor Sterrenkunde KU Leuven, Celestijnenlaan 200 B, B-3001 Leuven, Belgium

† European Southern Observatory, Casilla 19001, Santiago 19, Chile

‡ Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-5300 Bonn 1, Germany

§ Washburn Observatory, 475 North Charter Street, Madison, Wisconsin 53706, USA

A SMALL number of intrinsically luminous, low-mass stars have recently been shown to have an extremely peculiar elemental abundance pattern. Their photospheric carbon, nitrogen, oxygen and sulphur abundances are within an order of magnitude of solar values, but all other normally abundant metals are present in only trace amounts; in two stars, iron is deficient by nearly five orders of magnitude. Two possible explanations are that the low iron content is primordial, implying a very great age, whereas the CNO and S abundances have been acquired during evolution, or that the CNO and S abundances reflect the initial stellar composition and the low iron content is the result of chemical separation by dust formation. The latter hypothesis arises mainly because the abundance pattern of these stars is similar to that of interstellar gas¹, in which fractionation to dust plays an important part, but it is not easily understood how a process that must occur in the circumstellar envelope can so strikingly affect the photospheric abundances. Here we report the detection of appreciable amounts of zinc in the star HD52961, and argue that, because zinc will condense into dust only at rather low temperatures, its detection in near normal amounts is convincing evidence for the fractionation hypothesis.

The four stars BD+39°4926 (refs 2–4), HR4049 (refs 3, 5, 6), HD52961 (ref. 6) and HD44179 (refs 7, 8) are all intrinsically luminous, but their galactic latitudes suggest that they are of low mass. They must currently be in a short and advanced evolutionary stage, presumably between the asymptotic giant branch (AGB) and the central stars of planetary nebulae; this post-AGB interpretation is also consistent with the presence of an infrared excess, produced by dust heated by the star, in all but BD+39°4926. Their most remarkable common feature is their photospheric composition: they show almost solar CNO and S abundances (with respect to H), but very low Fe and, in one object or another, very low Ca, Al, Ti, Mg, Cr and Si abundances (when these elements can be measured). If the low metallic abundances are taken to represent the original composition of these stars, then the high CNO abundances may be attributed to nucleosynthesis in the stellar interiors followed by dredge-up to the surface. The fact that nearly solar amounts of CNO elements are produced is not necessarily a flaw in this hypothesis if one realizes that solar amounts of carbon have been produced in many carbon stars. But standard theory does not predict the high interior temperatures that would be necessary to produce sulphur in low-mass stars. Such a sulphur anomaly was first observed⁹ for the star HD46703, which has a similar but less extreme abundance pattern.

Following suggestions by Lambert *et al.*³ and by Venn and Lambert¹, Bond⁴ has proposed that the iron abundance in these stars is not primordial, but has been the result of chemical fractionation through dust formation. The main observational argument for this is the remarkable correlation between the photospheric abundances and the pattern of elemental depletions to dust in the interstellar medium (ISM). The elements that are not severely underabundant in these stars, CNO and S, happen to be those that are largely present in the gas phase in the ISM, whereas those missing from the stellar photospheres

are also highly depleted from the interstellar gas. The four stars would then still be interpreted as post-AGB objects, but with an initial composition not greatly different from the solar one, and the depleted elements being found in dust.

The main difficulty with Bond's hypothesis is in understanding how such chemical fractionation could occur. Stellar temperatures are too high for dust to form in the photosphere. In fact, observations of AGB stars indicate that dust forms at large distances from the stellar surface in these cool objects¹⁰, confirming the theoretical expectation that mass loss must precede dust formation.

A crucial test for deciding between the two hypotheses is the photospheric zinc abundance in these stars. This element shares with CNO and S the characteristic of being only slightly depleted in the interstellar gas^{1,11–13}. On the other hand, it belongs to the iron group, and so its abundance should follow that of iron if the underabundances are primordial. Detection of a high abundance of zinc is thus comprehensible only in the fractionation hypothesis. Unfortunately, we can expect to see reasonably strong zinc lines only in HD52961, the coolest of our four stars; these lines are the Zn I lines at 4,722.16 Å and 4,810.53 Å. As a matter of fact, the latter line was observed on a high-resolution ($R = 50,000$) spectrum of HD52961 which we obtained with the Coudé echelle spectrograph, equipped with a CCD detector and attached to the 1.4-m Coudé Auxiliary Telescope, at the European Southern Observatory (ESO) in January 1990; not being aware then of the importance of measuring zinc lines, we did not obtain a high-resolution observation of the former, somewhat stronger, line, but can report that it is clearly apparent on a low-resolution spectrum we obtained at ESO during February 1989. We have determined the zinc abundance for HD52961 from the equivalent width of the 4,810.53 Å line using the same methods as for analysis of the other elements⁶; the result is that zinc has an abundance of -1.3 with respect to the Sun in logarithmic units, and so is more than three orders of magnitude less depleted than iron. Furthermore, our high-resolution spectra do not so far reveal the presence of any other element that would have near-solar abundances in the photosphere of HD52961.

In Fig. 2a and b we show the observed deficiencies of elements in HD52961 and HD44179 as a function of their depletion in the ISM. As ISM depletions vary between different lines of sight and depend on density, we have averaged the values for various lines of sight^{11–14} and have used depletions for an ISM density of one H atom per cubic centimetre. For HD44179 the abundance of magnesium, an element that is moderately depleted in the ISM confirms the correlation seen for HD52961. We interpret Fig. 2 as strong evidence that the low iron abundances in

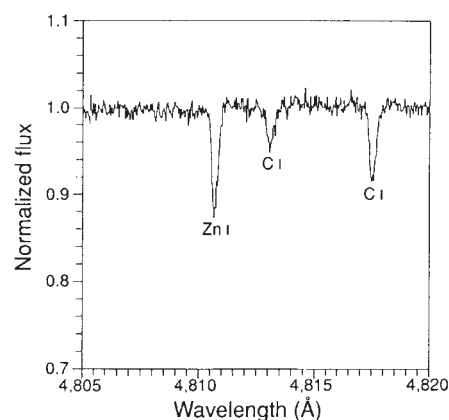


FIG. 1 High-resolution spectrum of HD52961 showing the Zn I line at 4,810.43 Å and two weak carbon lines. From the strength of the zinc line it is found that this element is underabundant by a factor of 20 in this star, whereas iron is underabundant by a factor of 65,000.

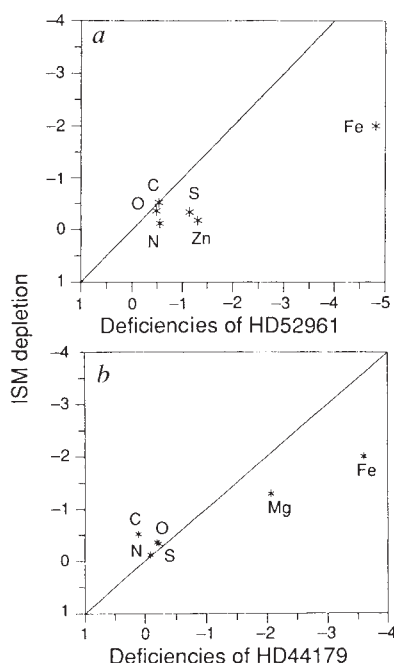


FIG. 2 *a*, Observed underabundances with respect to the solar value (in logarithmic units) of elements in HD52961 compared with the depletion of these elements in the interstellar gas. The five elements that are least depleted in the ISM seem not to be severely underabundant in HD52961; all other elements are, so much so that only for iron are spectral lines observed. The position of the observations relative to the bisector shows that the underabundance of iron in the star is a larger factor than the ISM depletion. *b*, As *a*, but for HD44179. This star is too hot for zinc to be observable, but allows the detection of magnesium, which is less depleted than iron in the ISM and shows an intermediate abundance in HD44179. Note the high carbon abundance, which is probably affected by evolution.

these stars are not primordial, but are a result of chemical separation.

It remains unclear how and where the separation operates. As stated earlier, dust formation in a stellar photosphere is unlikely, so the fractionation must have occurred in dust formed in matter expelled from the star and the gas must later have accreted again. This fractionation process must be extremely efficient, because the slopes in Fig. 2 show that the depletion in the atmospheres of these stars is larger than in the ISM, and because it seems to take place during a short evolutionary phase. A similar but less extreme abundance pattern occurs in the λ Bootis stars¹, which are main-sequence A stars with effective temperatures between 8,000 and 9,000 K. Accretion has been suggested as the explanation of the abundances of these stars^{1,15}; the analogy with our objects, and the presence of an infrared excess for some λ Boo stars, may suggest that the accreting material is from circumstellar rather than interstellar origin for the λ Boo stars as well.

A striking fact is that our stars occupy a very restricted zone in the Hertzsprung–Russell diagram. HR4049, HD44179 and BD +39° 4926 all have an effective temperature of $\sim 7,500$ K and a gravity of about $\log g = 1$ in c.g.s. units, and HD52961 is of similar luminosity but cooler, with an effective temperature of $\sim 6,000$ K; as mentioned above, a sulphur overabundance is also observed for HD46703, also at 6,000 K (ref. 9). The lower limit in temperature may be set by the condition that the stellar envelope should not be convective, because mixing would then rapidly wash out any superficial abundance anomalies. Not knowing of any hotter post-AGB stars with circumstellar dust near to the star, we cannot draw conclusions about any upper limit in temperature for the phenomenon, but the clustering of three stars at 7,500 K is suggestive.

The fractionation process being so efficient, it may have

implications for the interpretation of abundances of other metal-poor stars as well. It is well known that carbon and oxygen are overabundant with respect to iron in the most iron-poor halo stars¹⁶, and it has been reported that aluminium is underabundant with respect to iron in these stars¹⁷. These trends are usually interpreted in terms of nucleosynthetic processes in the earliest generations of supernovae, but our findings, and the fact that aluminium has a condensation temperature still higher than iron, suggest that fractionation may also have some role here. $\square$

Received 3 December 1991; accepted 3 March 1992.

- Venn, K. A. & Lambert, D. L. *Astrophys. J.* **363**, 234–244 (1990).
- Kodaira, K., Greenstein, J. L. & Oke, J. B. *Astrophys. J.* **159**, 485–512 (1970).
- Lambert, D. L., Hinkle, K. H. & Luck, R. E. *Astrophys. J.* **333**, 917–924 (1988).
- Bond, H. E. in *Evolution of Stars: The Photospheric Abundance Connection* (eds Michaud, G. & Tutukov, A.) 341–346 (Kluwer, Dordrecht, 1991).
- Lamers, H. J. G. L. M. et al. *Astr. Astrophys.* **154**, L20–L23 (1986).
- Waelkens, C. et al. *Astr. Astrophys.* **251**, 495–504 (1991).
- Cohen, M. et al. *Astrophys. J.* **196**, 179–189 (1976).
- Waelkens, C. et al. *Astr. Astrophys.* **256**, L15–L18 (1992).
- Bond, H. E. & Luck, R. E. *Astrophys. J.* **312**, 203–212 (1987).
- Tielens, A. G. G. M. in *From Miras to Planetary Nebulae: Which Path for Stellar Evolution?* (eds Mennessier, M. O. & Omont, A.), 186–200 (Editions Frontières, Paris, 1990).
- Harris, A. W. & Mas Hesse, J. M. *Mon. Not. R. astr. Soc.* **220**, 271 (1986).
- Van Steenberg, M. E. & Shull, J. M. *Astrophys. J.* **330**, 942–963 (1988).
- Jenkins, E. B. in *Interstellar Processes* (eds Hollenbach, D. J. & Thronson, H. A. Jr) 533–560 (Astrophysics and Space Science Library, Reidel, Dordrecht, 1987).
- Phillips, A. P., Gondhalekar, P. M. & Pettini, M. *Mon. Not. R. astr. Soc.* **200**, 687 (1982).
- Charbonneau, P. *Astrophys. J.* **372**, L33–L36 (1991).
- Gustafsson, B. in *ESO Workshop on Stellar Evolution and Dynamics in the Outer Halo of the Galaxy* (eds Azzopardi M & Mateucci, F.) 33–45 (1987).
- Magain, P. *Astr. Astrophys.* **179**, 176–180 (1987).

ACKNOWLEDGEMENTS. Based on observations collected at the European Southern Observatory, La Silla, Chile.

Energy transfer and vibrational effects in the dissociation and scattering of D_2 from Cu(111)

A. Hodgson, J. Moryl, P. Traversaro & H. Zhao

Surface Science Research Centre and Department of Chemistry, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

ACTIVATED surface dissociation of adsorbed molecules is of basic importance to surface chemistry, but in only a few systems are details of the molecule–surface potential-energy surface well understood. Central to this problem are the processes of energy disposal and transfer during the molecule–surface collision and the manner in which molecular motions (translational, rotational and vibrational) are channelled into the dissociative coordinate. The dissociation of hydrogen and deuterium molecules on a copper surface provides a prototypical system for studying these processes. Here we describe measurements of the scattering of D_2 from a Cu(111) surface as a function of incident translational energy, angle and vibrational state of the incoming molecules. Molecules in the ground and first excited vibrational levels show very different translational-energy thresholds for removal (0.6 and 0.35 eV respectively). At higher incident energies, efficient transfer of translational to vibrational energy occurs in those ground-state D_2 molecules that survive the collision without dissociating. These data will allow a better test of the various potential-energy curves that have been proposed for this system in the dissociative region.

The dissociation of a small molecule is often an activated process, requiring considerable energy to be placed in the molecule before it can dissociate to form chemisorbed atomic or radical products. Many studies have investigated the influence of factors such as surface temperature and translational energy on dissociation probability. To understand the detailed dynamics on the femtosecond timescale of the dissociation process, it is necessary to compare experimental measurements with model scattering calculations on an empirical or *ab initio*

potential energy surface (PES). Measurements of energy transfer and scattering, under conditions where dissociation occurs, provide insight into the region of the PES near the barrier to dissociation, and thus into the dissociation process itself.

It has often been assumed that the barrier to dissociative chemisorption lies in the molecule-surface coordinate and that the dissociation is therefore promoted by translational energy perpendicular to the surface¹. For the H_2 , D_2 /Cu system there is evidence, both from state-resolved measurements following recombinative desorption² and from molecular-beam measurements³ of the dissociative chemisorption probability, S_0 , for a molecule striking a clean surface, that the vibrational motion of the gas molecules is critical in this dissociation. Kubiak *et al.*² showed that recombination preferentially generates vibrationally excited molecules, indicating (using the principle of microscopic reversibility) that dissociation is favoured for vibrationally excited molecules. Hayden and Lamont³ showed that the probability of dissociative chemisorption depends not only on the translational energy of the H_2 or D_2 molecules but also on their internal energy. This agrees with the prediction⁴, based on a theoretical H_2 -Cu₂ potential energy surface (PES), that the barrier to dissociation will be late in the reaction coordinate, at extended H_2 bond separations. Scattering calculations of the sticking probability S_0 for various empirical and theoretical PES⁵⁻⁹ predict a translational energy threshold strongly dependent on vibrational state, above which sticking becomes efficient. Rettner *et al.*¹⁰ have recently extended previous work^{11,12} on Cu(111) parameterizing and fitting sticking functions for the different vibrational states to adsorption data.

Vibrationally excited $H_2(v=1)$ is preferentially removed at a Cu(111) surface, and this is consistent with efficient dissociative chemisorption of the vibrationally excited state¹³. Here we report the direct observation of the translational-energy and angle dependence for the removal of individual vibrational states in D_2 -Cu(111) collisions, and efficient vibrational excitation of the reflected D_2 beam as the threshold for dissociative chemisorption of ground-state $D_2(v=0)$ is reached. We have monitored the reflectivity of D_2 in the $v=0$ and $v=1$ states as a function of the incident translational energy and scattering angle. Experimental details have been described elsewhere¹³. A translationally and vibrationally hot beam of D_2 (up to 1,700 K) was scattered from a Cu(111) surface. The translational energy distribution was selected using seeding techniques and measured from the

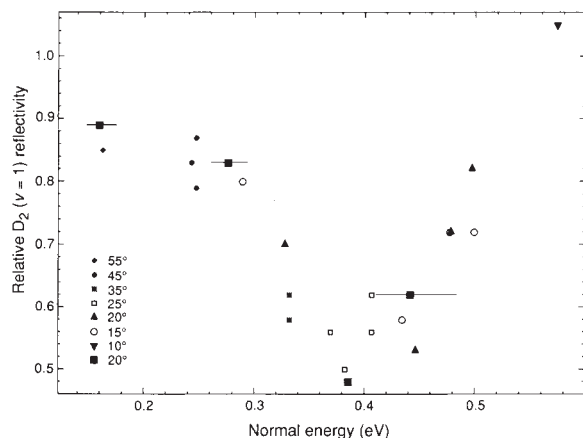


FIG. 1 Relative reflectivity, $R_{10}(E_{\perp})$, for $D_2(v=1, J=4)$ and $D_2(v=0, J=4)$ scattered from a Cu(111) surface, plotted against the mean translational energy perpendicular to the surface. $R_{10}(E_{\perp})$ is insensitive to the angular resolution used (5–30°) and reflects the difference in probability of a particular vibrational level being created or removed during the surface collision. The points (■) show data taken using the chopper to select a narrow velocity group. The abrupt drop in the relative reflectivity of $D_2(v=1)$ near 0.35 eV is attributed to the onset of dissociative chemisorption and vibrational relaxation for this state.

time of flight (TOF) of the incident molecules in the usual way¹⁴. The high nozzle temperatures necessary to provide the required energies give translational energy distributions which, as for other studies of the sticking coefficient^{3,11,12}, are rather broad. Surface cleanliness was ensured by maintaining the crystal above the desorption temperature for D_2 . The population of individual rotational and vibrational levels in the incident and scattered beams was measured using resonance-enhanced multiphoton ionization through the $E, F^1\Sigma_g^+$ state. The reflectivity of $D_2(v=1, J)$ was then compared to $(v=0, J)$ by measuring the populations of the two levels both in the incident beam and in the flux scattered off the surface. This gives $R_{10}(E_T)$, the reflectivity of the $v=1$ state relative to that of $v=0$, as a function of incident

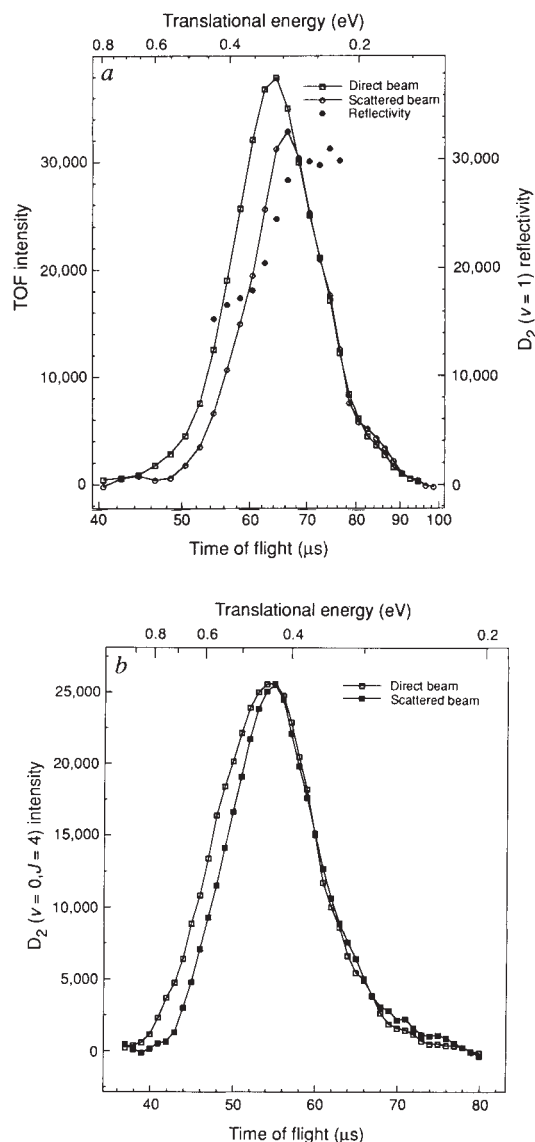


FIG. 2 Time-of-flight (TOF) profiles of (a) $D_2(v=1, J=4)$ and (b) $D_2(v=0, J=4)$ scattered from a Cu(111) surface, compared with the incident TOF distributions. *a*, Loss of $D_2(v=1, J=4)$ at energies above 0.3 eV. The scattered $D_2(v=1)$ signal has been normalized to the incident TOF using the measured reflectivities of the two state. If convolution effects from the narrow (4-μs) chopper gate are ignored, the energy is related directly to the TOF, as shown on the top scale. The relative reflectivity $R_{10}(E_{\perp})$ can be obtained directly and is shown by the solid symbols; in agreement with Fig. 1, it falls rapidly near 0.35 eV. At these energies the TOF profiles for the direct and scattered $D_2(v=1, J=4)$ are identical. *b*, $D_2(v=0, J=4)$ reflectivity at higher energies, the scattered signal having been arbitrarily scaled to the direct beam. The reflectivity drops steadily above ~0.5 eV, consistent with the onset of dissociative chemisorption for $D_2(v=0)$.

translational energy E_T and angle θ_i . In some of these experiments the energy resolution was improved by using a narrow (4- μ s) chopper window and measuring the reflectivity as a function of flight time.

The relative reflectivity of $D_2(v=1, J=4)$ scattered from Cu(111) is shown in Fig. 1 as a function of the incident translational energy perpendicular to the surface, $E_\perp$. $R_{10}(E_\perp)$ is ~ 0.9 at low energies but falls near 0.35 eV to ~ 0.5 . This is caused by an abrupt increase in the rate of removal of $D_2(v=1)$ and is attributed to the onset of dissociative chemisorption and vibrational relaxation of the ($v=1$) level, consistent with the barrier inferred for this level from measurements of the sticking coefficient¹⁰. Figure 2 shows the time of flight of $D_2(v=1, J=4)$ incident and scattered from the surface with $E_\perp \approx 0.3$ eV. The flight time from the crystal to the detection region is negligible and, as expected, the incident and scattered $D_2(v=0)$ TOF profiles and the incident ($v=1$) TOF profile are identical. The angular distributions of the scattered molecules are also the same, and differences in the ($v=1$) TOF profile reflect gain or loss of $D_2(v=1, J=4)$ flux as a function of the incident translational energy. The relative reflectivity of the $v=1$ state again shows a sharp reduction near 0.35 eV as sticking becomes efficient. The removal of $D_2(v=1)$ follows a normal ($\cos^2 \theta_i$) energy scaling, and R_1 can be fitted reasonably well to an expression of the form¹⁵ $R_1 = \Lambda/2 \{1 + \tanh[(E_\perp - E_v)/W]\}$ where E_v , Λ and W are parameters giving the energy threshold (0.35 eV), width (0.15 eV) and limiting removal probability (0.5), respectively. This behaviour is in reasonable agreement with the threshold for $v=1$ dissociative chemisorption obtained from sticking measurements¹⁰ although the quantity measured here includes all loss channels for $D_2(v=1)$, including vibrational relaxation.

At higher energies (Fig. 1), R_{10} increases until it becomes greater than 1. This is attributed to two processes, sticking of the ground vibrational level and efficient vibrational excitation ($1 \leftarrow 0$) of those ground-state molecules which do not stick. By comparing $D_2(v=0)$ TOF profiles for the scattered and incident beams at an incident energy of ~ 0.6 eV, we can observe the attenuation of the scattered ($v=0$) flux. The onset of removal of $D_2(v=0, J=4)$ occurs near 0.5 eV and increases rather slowly with energy compared with the $v=1$ removal, reaching a value

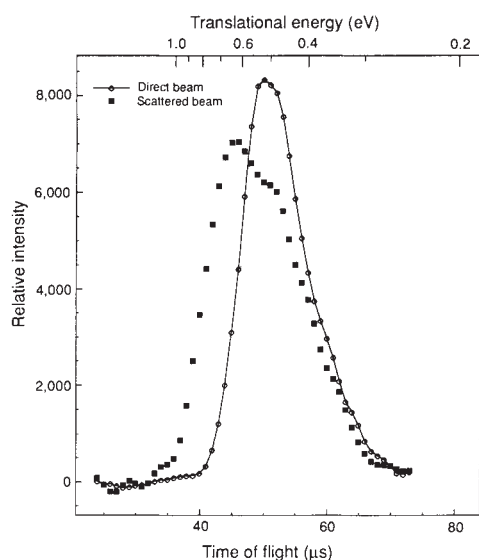


FIG. 3 Time-of-flight profiles for $D_2(v=1, J=4)$ in the direct beam ($\circ$), and scattered from the surface ($\blacksquare$) ($\theta_i = 20^\circ$, $\theta_f = 20^\circ$). The scattered flux results from a competition between removal by dissociative chemisorption or relaxation above 0.35 eV, and vibrational excitation of $D_2(v=0)$ which becomes efficient above $E_\perp = 0.6$ eV. The large enhancement in the scattered flux at high energies reflects the high probability for T $\rightarrow$ V transfer.

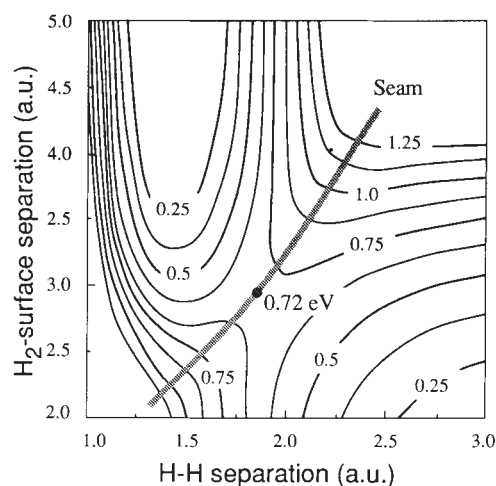


FIG. 4 Two-dimensional representation of the potential energy surface (PES) for the D_2 -Cu system⁸. The barrier or seam occurs at extended D_2 separations and activation by translation must occur through conversion of motion along the Cu- D_2 coordinate into a stretching of the D_2 bond (T $\rightarrow$ V coupling). Those molecules that fail to cross the barrier and dissociate may return to the gas phase vibrationally excited. By contrast a vibrationally excited molecule approaching the barrier already has energy in the D_2 stretching coordinate¹⁸ and will dissociate efficiently with a lower translational barrier. Axes are in atomic units (a.u.) where 1 a.u. = 0.529 Å.

of $(60 \pm 30)\%$ by 0.7 eV. As the incident energy is increased, the flux of D_2 scattered from the surface in the $v=1$ state exceeds that which was incident (Fig. 3): vibrational excitation ($1 \leftarrow 0$) of the ground state occurs with remarkable efficiency, despite the size of the vibrational quanta involved (0.36 eV), and more than compensates for the removal of $v=1$ flux by dissociation and relaxation. The vibrational excitation is consistent with an $E_\perp$ energy scaling, rather than E_T , increasing from a threshold near $E_\perp = 0.6$ eV. Assuming that the incident $v=1$ population is given by the nozzle temperature, the probability for vibrational excitation reaches $(40 \pm 30)\%$ at $E_\perp = 0.8$ eV.

The only possible source of the 0.36-eV energy required for D_2 vibrational excitation is the translational energy of the incident molecules. The normal energy scaling indicates that only the component of the translational energy perpendicular to the surface contributes to vibrational excitation. Such efficient, direct translational to vibrational (T $\rightarrow$ V) energy transfer has not previously been observed in surface scattering; it indicates coupling of these motions in the region of the PES near to the barrier for dissociation. Vibrational excitation has been observed¹⁶ for NO scattering from Ag(111) but that excitation was strongly dependent on surface temperature, rather than on the translational energy, and was attributed to electronic excitation through electron-hole pairs. In contrast, vibrational excitation of D_2 from Cu(111) was insensitive to surface temperature. The angular distribution of the $D_2(v=1)$ produced by the T $\rightarrow$ V transfer was measured and peaks close to the specular direction, rather than along the surface normal. This is consistent with direct collisional T $\rightarrow$ V excitation during a repulsive surface collision, rather than with excitation through a long-lived molecular precursor with an extended D_2 bond¹⁷. Saturation of the surface with D_2 prevents dissociation and was found to inhibit vibrational excitation, showing the importance of the interaction potential with the clean metal surface for vibrational excitation.

The barrier to dissociative chemisorption of D_2 on Cu(111) occurs at extended D_2 bond lengths^{4,8} (see Fig. 4), and sticking measurements^{3,11,12} and calculations^{5,7,8} have shown the importance of vibrational energy for dissociation. Near the barrier, the potential strongly couples the vibrational (dissociative) and translational motion, converting translational energy into

motion along the D_2 stretching coordinate and giving rise to vibrational excitation or dissociation. For $D_2/Cu(111)$, vibrational excitation starts to occur at energies just above the threshold for $D_2(v=0)$ dissociation, indicating that it is intimately associated with lengthening of the D_2 bond in the region near the transition state for dissociation. The high efficiency for both vibrational excitation and $D_2(v=1)$ removal indicates that a large fraction of the molecules enter the region of the PES responsible for dissociation and $T \rightarrow V$ coupling, suggesting that the steric requirements on D_2 bond angle and impact site are not severe and that low-dimensional scattering calculations may be a reasonable approach^{5,7,8}. We expect that modelling of the $T \rightarrow V$ transfer, as well as the sticking coefficient S_0 , will provide a detailed picture of the PES for this system in the region of the barrier to dissociation. □

Received 16 December 1991; accepted 6 March 1992.

1. Lennard-Jones, J. E. *Trans. Faraday Soc.* **28**, 333–359 (1932).
2. Kubiak, G. D., Sitz, G. O. & Zare, R. N. *J. chem. Phys.* **83**, 2538–2551 (1985).
3. Hayden, B. E. & Lamont, C. L. A. *Phys. Rev. Lett.* **63**, 1823–1826 (1989).
4. Harris, J. & Anderson, S. *Phys. Rev. Lett.* **55**, 1583–1586 (1985).
5. Holloway, S., Halstead, D. & Hodgson, A. *J. Electron Spectrosc.* **45**, 207–213 (1987).
6. Harris, J. *Surf. Sci.* **221**, 335–345 (1989).
7. Hand, M. R. & Harris, J. *J. chem. Phys.* **92**, 7610–7617 (1990).
8. Hand, M. R. & Holloway, S. *J. chem. Phys.* **91**, 7209–7219 (1989).
9. Nielsen, U., Halstead, D., Holloway, S. & Nørskov, J. K. *J. chem. Phys.* **93**, 2879–2884 (1990).
10. Rettner, C. T., Auerbach, D. J. & Michelsen, H. A. *Phys. Rev. Lett.* **68**, 1164–1167 (1992).
11. Anger, G., Winkler, A. & Rendulic, K. D. *Surf. Sci.* **220**, 1–17 (1989).
12. Berger, H. F., Leisch, M., Winkler, A. & Rendulic, K. D. *Chem. Phys. Lett.* **175**, 425–428 (1990).
13. Hodgson, A., Moryl, J. & Zhao, H. *Chem. Phys. Lett.* **182**, 152–158 (1991).
14. Auerbach, D. J. in *Atomic and Molecular Beam Methods* (ed. Scoles, G.) (Oxford University Press, 1988).
15. Michelsen, H. A. & Auerbach, D. J. *J. chem. Phys.* **94**, 7502–7520 (1991).
16. Rettner, C. T., Kimman, J., Fabre, F. & Auerbach, D. J. *Surf. Sci.* **192**, 107–130 (1987).
17. Gadzuk, J. W. & Holloway, S. *J. chem. Phys.* **82**, 5203–5215 (1985).
18. Polanyi, J. C. & Wong, W. H. *J. chem. Phys.* **51**, 1439–1450 (1969).

ACKNOWLEDGEMENTS. We thank S. Holloway for discussion. This work was supported by the SERC.

Cation microsegregation and ionic mobility in mixed alkali glasses

B. Vessal*, G. N. Greaves*†, P. T. Marten*,
A. V. Chadwick‡, R. Mole‡ & S. Houde-Walter§

* Department of Chemistry, Keele University, Staffordshire ST5 5BG, UK

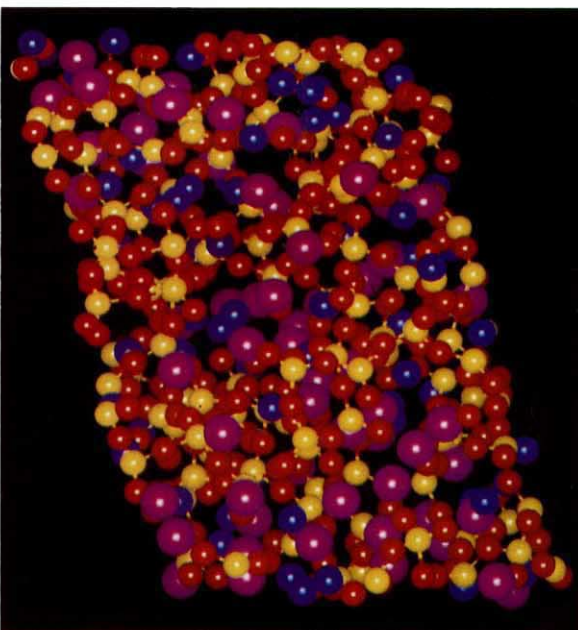
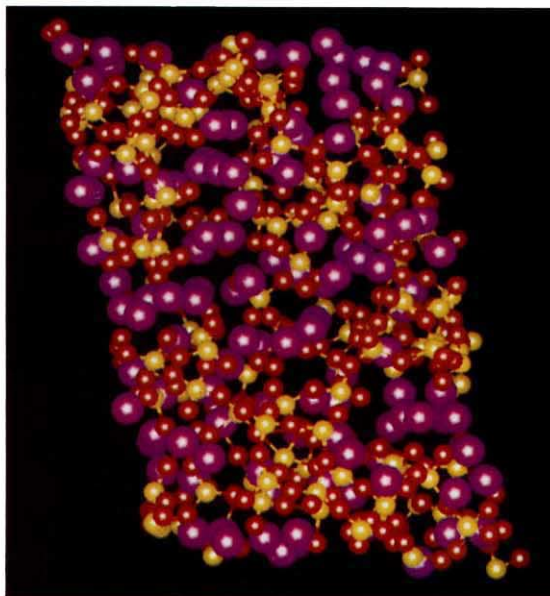
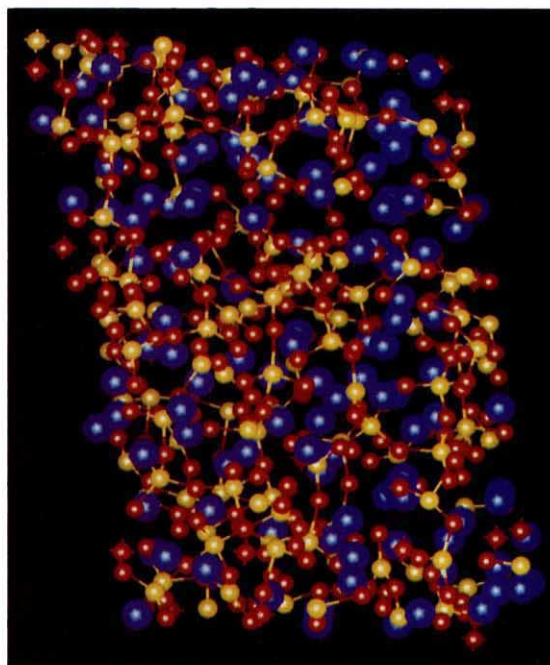
† SERC Daresbury Laboratory, Warrington, Cheshire WA4 4AD, UK

‡ University Chemical Laboratory, University of Kent, Canterbury CT2 7NH, UK

§ Institute of Optics, University of Rochester, New York 14627, USA

MUCH is known about short-range ($<5 \text{ \AA}$) structural order in oxide glasses from experimental probes of local structure such as X-ray absorption fine structure (XAFS)¹, but over the medium range (5–20 Å) their structures are poorly understood. Computer simulations based on measured parameters for local atomic environments, however, can provide structural models on the nanometre scale, which enable dynamic properties such as ionic transport to be considered. Here we describe a molecular dynamics simulation of the effects of mixed alkali cations on the structure of binary silicate glasses. It is well known that the ionic conductivity of alkali glasses falls markedly when more than one alkali is present². We demonstrate that the alkalis segregate from the silicate network over distances of a few ångströms. Although ionic mobility is expected to be higher in these microsegregated regions than in the surrounding silicate network, we suggest that stochastic mixing of alkalis nevertheless impedes the hopping process of a given alkali ion. This is manifest as an increase in the average activation energy for hopping and results in a lowering of the total ionic conductivity.

Continuous random network (CRN) models can readily reproduce the diffraction data of silica^{3–5}, and molecular dynamics (MD) simulations have modelled its structure by



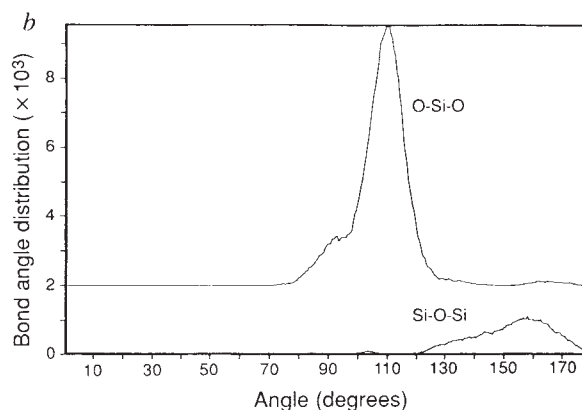
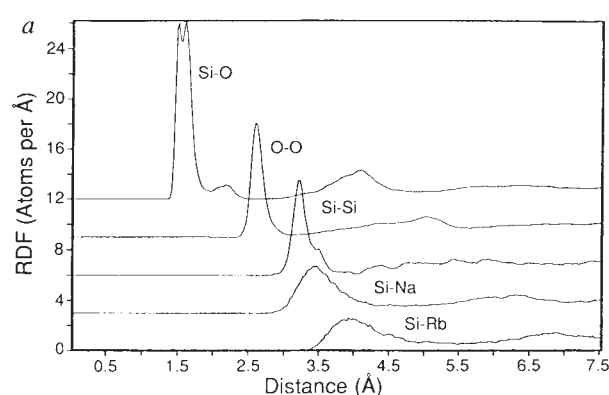


FIG. 2 Atomic and bond-angle distributions for MD simulated glasses. *a*, Partial RDFs (atoms per Å) for Si-O, O-O, Si-Si, Si-Na and Si-Rb in the NaRbSi₂O₅ structure shown in Fig. 1. *b*, Distribution (configurations per

degree) of O-Si-O (upper) and Si-O-Si (lower) bond angles for the Rb₂Si₂O₅ glass in Fig. 1.

following the liquid-glass transition^{6,7}. Silicate glasses have received less attention^{8,9}, partly because of their added complexity and partly because of the paucity of structural information available from diffraction experiments. XAFS¹ and solid-state NMR¹⁰ have revealed that the immediate local environments of glass-modifying cations such as the alkali metals are well defined, as are the configurations of SiO₄ units present for a particular composition. On the basis of these results, a modified random network (MRN) was proposed in which all atom types—alkalis, silicons, and bridging and non-bridging oxygens—were fully coordinated¹¹. The model predicted segregation on the atomic scale of the directionally bonded glass formers from the ionically packed glass-modifying components. These micro-inhomogeneities are the analogues of the molecular and intermolecular units, respectively, in the layer and chain structures that characterize the crystalline silicates.

The model potentials used here have been described previously^{12,13} and used successfully in recent MD simulations^{7,9}. The present simulations were done at constant volume in a monoclinic box with periodic boundary conditions. The compositions varied from Na₂Si₂O₅ to Rb₂Si₂O₅, including a range of mixed alkali compositions. Each crystalline structure was first melted at 6,000 K for 7.5 ps (2,500 time steps) and then cooled successively for the same duration at 2,500, 1,200, 625 and 293 K to obtain the annealed glass structure. Examples of annealed structures for single and mixed alkali compositions are shown in Fig. 1. The clustering of cations is evident in the medium-range structure of all three glasses.

The partial radial distribution functions (RDFs) for silicon and oxygen in the simulated mixed alkali glass NaRbSi₂O₅ are shown in Fig. 2*a* and are broadly typical of all the glasses studied. The sharpest features belong to the principal intra-network correlations: Si-O, O-O and Si-Si. They are similar to the equivalent features found in modelling silica⁷, except that the Si-O peak at 1.61 Å is split into two contributions at 1.51 and 1.62 Å. Similar splits occur in the crystalline lithium and sodium disilicates. Compared with α -quartz, in these layered structures the bridging oxygen distance is increased and the non-bridging oxygen distance decreased, distorting the SiO₄ tetrahedra and increasing the mean Si-O distance. The splitting of Si-O correlations for the simulated glass structure in Fig. 2*a* decreases with sodium content, disappearing for Na₂Si₂O₅ glass. We have found similar behaviour in simulations of (K, Cs)₂Si₂O₅ glasses, where Si-O distances are split for Cs-rich compositions but not for K₂Si₂O₅ glass (results not shown). Recent XAFS measurements for silicon in single alkali silicate glasses reveal commensurate increases in the average Si-O distances¹⁴.

Figure 2*b* shows typical bond-angle distributions for the network atoms in the simulated structures. The O-Si-O angles are constrained by the three-body term to the tetrahedral angle of 109.5°. The distribution of Si-O-Si bond angles, which were not constrained, is much broader and asymmetric with a maximum close to 155° and is similar to that previously obtained by MD for silica⁷. It is clear from Fig. 2*a* and *b* that despite the depolymerizing effect of alkalis and non-bridging oxygens, the short-range geometry of the network component in modified glasses seems to be close to that of silica.

The partial RDFs for sodium and rubidium in NaRbSi₂O₅ glass (Fig. 3) are much broader than the network distributions shown in Fig. 2*a* and reflect the weaker alkali-oxygen bonding. The Na-O correlations peak at 2.4 Å and Rb-O correlations at 3.0 Å. In the total RDF obtained using diffraction techniques, these distances are obscured by the strong O-O peak at 2.6 Å and can only be accurately measured by XAFS. In earlier experiments on Na₂Si₂O₅ glass, we recorded a Na-O separation of 2.3 Å¹⁵ and we recently measured a Rb-O distance of 2.8 Å in Rb₂Si₂O₅ glass (unpublished data). The alkali-oxygen distances obtained here are therefore reasonably close to experiment and reflect the accuracy of the potentials used for sodium and rubidium¹³.

In contrast to silicon, the coordination numbers of alkalis obtained in the simulated structures are significantly greater than 4, as found earlier⁸. The distribution of alkali coordination numbers peaks between 5 and 6 for all the MD glasses studied here and as such agrees with the results of Na, K, Rb and Cs XAFS^{14,15}. Moreover, the M-O-M/Si and O-M-O bond-angle distributions (M is Na or Rb) peak at smaller angles than the corresponding Si-O-Si and O-Si-O distributions that

FIG. 1 Molecular graphics representations of the simulated glasses: Na₂Si₂O₅ (top) Rb₂Si₂O₅ (middle) and NaRbSi₂O₅ (bottom). Periodic boundary conditions and monoclinic boxes containing 576 atoms were used. Crystalline Na₂Si₂O₅ was the initial structure chosen with a box volume of 7,548.9 Å³, and this was increased as more rubidium was added. The long-range Coulomb contribution was calculated using Ewald's method with full ionic charges on both the network and modifying atoms. We used four-range Buckingham potentials (one for each of Si-O, O-O and so on) to model the short-range interactions between network atoms, and simple Buckingham potentials for those between the modifier atoms and the oxygens. To simulate the directional bonding of the network, three-body interactions were used to replicate the tetrahedral O-Si-O linkage. This model potential has been parameterized against the elastic and dielectric constants of α -quartz, which are accurately reproduced⁷. In the annealed structures illustrated here, silicon atoms are represented by yellow, oxygen atoms by red, sodium by blue and rubidium by purple spheres, respectively. The silicon and oxygen atoms are shown smaller in the top two structures so that the depth dimension is clearer. The occasional atoms apparently floating at the surface of the monoclinic box are a consequence of periodic boundary conditions, their coordinated sites being translated by a lattice vector.

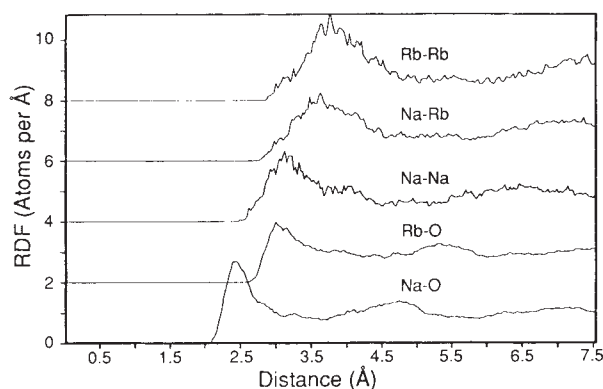


FIG. 3 Partial RDFs (atoms per Å) for Rb-Rb, Na-Rb, Na-Na, Rb-O and Na-O in simulated NaRbSi₂O₅ glass shown in Fig. 1.

characterize the network component. The maxima in the distributions of Rb-Si and Na-Si correlations, for instance, correspond to angles of $\sim 100^\circ$ and $\sim 95^\circ$, respectively. These angles suggest that the oxygens connecting alkalis to silicons (non-bridging oxygens) have a coordination number substantially higher than the doubly coordinated bridging oxygens connecting pairs of silicons in the network. The average coordination number of all types of oxygen in the simulated structures is ~ 3 . Of these, $\sim 60\%$ are bridging oxygens with a coordination number of 2 leaving the remaining non-bridging oxygens with a coordination number of ~ 5 . Thus the non-bridging oxygens have coordination numbers matching those of the alkalis, as has previously been inferred from XAFS measurements¹⁴.

Two further important results are contained in Fig. 3. Note first how the Na-O partial RDF exhibits, in addition to the nearest-neighbour peak at 2.4 Å, a further peak at 4.8 Å. Even though the concentration of Na₂O is only 16%, there are clearly Na-Na correlations at 3.1 Å and near 6.5 Å, exceeding what would be expected if sodium were evenly distributed throughout

the structure. Similarly, the Rb-O partial RDF contains a feature at 5.4 Å as well as that due to bonded oxygen at 3.0 Å. The distribution of Rb-Rb distances peaks at 7.5 Å and also at 3.8 Å. The identification of first (O), second (M), third (O) and fourth (M) nearest neighbours around every alkali demonstrates that these modifying oxides, although they are minor components in the glass, do indeed form clusters within the simulated networks as envisaged in the MRN model¹¹. This can be seen in Fig. 1 both for single and mixed alkali M₂Si₂O₅ glasses. The genesis of alkali-rich regions has been explored in sodium silicate glasses with MD techniques⁹. Our results establish that microsegregation is ubiquitous in disilicate glasses, irrespective of the alkali composition. Geometrically it is a direct result of the high coordination numbers of alkalis and of the oxygens to which they are bound, compared with the remaining silicons and bridging oxygens.

Figure 3 also demonstrates that when two alkalis are present in the same glass structure, they are intimately mixed within the segregated regions. The Na-Rb partial RDF has equivalent features to the Na-Na and Rb-Rb distributions. The intensities are of similar magnitude, and the first and second peaks fall roughly midway between the corresponding peaks for the single alkali pairs, demonstrating that within the modified regions of the glass structure the alkalis are randomly distributed. Stochastic mixing of sodium and rubidium is evident in Fig. 1 in the alkali-rich portions of the NaRbSi₂O₅ glass structure. In a recent XAFS study of KCsSi₂O₅ glasses the juxtaposition of the two alkalis was inferred from changes in their Debye-Waller factors when mixing occurred, the order around potassium improving at the expense of that around caesium¹⁴. Similar findings have been reported for changes in far-infrared bands in alkali borate glasses¹⁶. There is a small increase in the nearest-neighbour disorder around rubidium as sodium is added to the composition. The opposite is true for the oxygen shell of sodium, which becomes slightly better ordered if rubidium is added.

The mixed alkali effect² is primarily related to an increase, ΔW , in the activation energy W of electrical conductivity σ . Figure 4 shows the temperature dependence of the d.c. conductivity, $\sigma T = (\sigma T)_0 e^{-W/kT}$, measured¹⁷ for single and mixed-alkali Na-Rb silicate glasses similar in composition to the simulated structures we have been describing. Clearly, although all three curves converge on similar values of $(\sigma T)_0 \approx 10^5 \Omega^{-1} \text{cm}^{-1}$ (see ref. 14), the activation energies are different. In particular, $W_{\text{Na}} < W_{\text{Rb}} < W_{\text{Na+Rb}}$, and the increase as a result of mixing, ΔW , is 0.36 eV. If the size and binding energy of the two alkalis are different this general behaviour can be predicted both from a simple statistical model¹⁸ and from percolation arguments¹⁹. In either case ionic transport is sustained by nearest-neighbour hopping, which will be hindered by the intimate mixing of alkalis that we observe in these simulations of glass structure. □

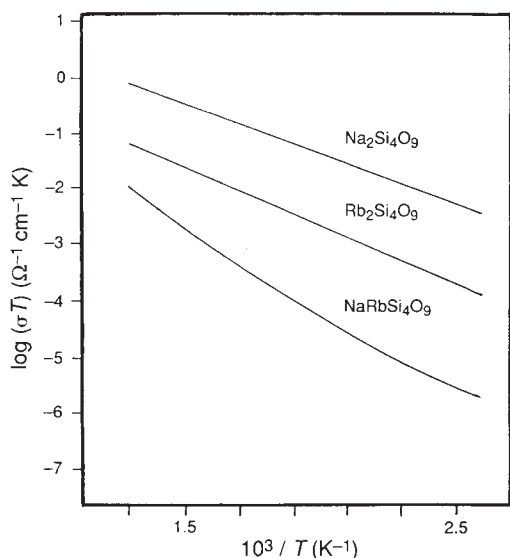


FIG. 4 Temperature dependence of d.c. electrical conductivity measured in M₂Si₄O₉ glasses: Na₂Si₄O₉ (top), Rb₂Si₄O₉ (middle) and NaRbSi₄O₉ (bottom). Activation energies, W , are as follows: 0.73 eV (W_{Na}), 0.84 eV (W_{Rb}) and 1.14 eV ($W_{\text{Na+Rb}}$). The average activation energy of the two single-alkali glasses, $(W_{\text{Na}} + W_{\text{Rb}})/2$, is 0.78 eV, and so ΔW , the increase due to mixing, is 0.36 eV.

Received 18 December 1991; accepted 27 February 1992.

- Greaves, G. N., in *Glass Science and Technology*, Vol. 4B (eds Uhlmann, D. R. & Kreidl, N.) 1-76 (Academic, London, 1990).
- Day, D. E. *J. non-cryst. Solids* **21**, 343-372 (1976).
- Bell, R. J. & Dean, P. *Phil. Mag.* **25**, 1381-1398 (1972).
- Gladden, L. F. *J. non-cryst. Solids* **119**, 318-330 (1990).
- Keen, D. A. & McGreevy, R. L. *Nature* **344**, 423-425 (1990).
- Mitra, S. K. *Phil. Mag.* **B45**, 529-548 (1982).
- Vessal, B., Amini, M., Fincham, D. & Catlow, C. R. A. *Phil. Mag.* **B60**, 753-775 (1989).
- Soules, T. F. *J. chem. Phys.* **71**, 4570-4558 (1979).
- Huang, C. & Cormack, A. N. *J. chem. Phys.* **93**, 8180-8186 (1990).
- Dupree, R., Holland, D., McMillan, P. W. & Pettifer, R. F. D. *J. non-cryst. Solids* **68**, 399-410 (1984).
- Greaves, G. N. *J. non-cryst. Solids* **71**, 203-217 (1985).
- Vessal, B., Leslie, M. & Catlow, C. R. A. *Molec. Simul.* **3**, 123-136 (1989).
- Parker, S. C. thesis, Univ. of London (1982).
- Greaves, G. N. et al. *Phil. Mag.* **A64**, 1059-1072 (1991).
- Greaves, G. N., Fontaine, A., Raoux, D. & Gorman, S. J. *Nature* **293**, 611-616 (1981).
- Kamitsos, E. I., Patsis, A. P. & Chrysoskos, G. D. *Phys. Chem. Glasses* **32**, 219-221 (1991).
- Mole, R. thesis, Univ. of Kent (1991).
- Catlow, C. R. A. *Phys. Stat. Sol.* **46**, 191-198 (1978).
- Bunde, A., Maas, P. & Ingram, M. D. *Ber. Bunsenges. Phys. Chem.* **95**, 977-981 (1991).

ACKNOWLEDGEMENTS. We thank R. G. Bell, C. R. A. Catlow and C. Freeman for discussions and assistance with molecular graphics.

Effect of temperature on shock metamorphism of single-crystal quartz

Falko Langenhorst, Alexander Deutsch, Dieter Stöffler & Ulrich Hornemann*

Institut für Planetologie, Wilhelm-Klemm-Straße 10, D-4400 Münster, Germany

* Ernst-Mach-Institut, Hauptstrasse 18, D-7858 Weil am Rhein, Germany

FEATURES characteristic of shock metamorphism in target rocks are the main diagnostic tool for recognizing impact phenomena on the Earth and other planetary bodies¹⁻⁴, and experimentally calibrated shock effects in silicate minerals have been important in elucidating the pressure histories of these rocks. Except for a few preliminary results for experimentally shocked pre-heated quartzites^{5,6}, all available calibration data are based on experiments performed with the targets at room temperature⁷⁻¹¹. Observations at Vredefort^{12,13} and at the Sudbury impact structure¹⁴⁻¹⁶ indicate, however, that considerable shock stresses occur in deep-seated crustal rocks which are at elevated temperatures during large cratering events. High-temperature shock metamorphism must also have been of great importance in the collision history of meteorite parent bodies in the early Solar System. Here we report the results of shock experiments on single-crystal quartz heated to 630 °C, which show that the physical properties of shocked quartz (refractive indices, lattice parameters, amorphization, type and frequency of planar deformation features) depend strongly on the pre-shock temperature. We conclude that existing shock-wave barometers^{1,8,9,11} cannot be applied to high-temperature target rocks, and that new barometers independent of pre-shock temperature will be required to understand the shock pressure history of terrestrial and planetary impact formations.

Hypervelocity collisions of still-molten or at least 'hot' planetary objects were common in the early Solar System, as indicated by observations of meteorites^{3,4}. On Earth, the Sudbury structure with a transient cavity reaching the lower crust^{15,16} is an instructive example of an impact on hot target rocks. Assuming a normal geothermal gradient, one can predict a temperature of 600–700 °C for rocks at a depth of ≤ 25 km. Another probable case of high-temperature impact metamorphism is at Vredefort, where there is a circular structure for which

a superposition of shock metamorphism on static metamorphism at temperatures of up to 900 °C has been proposed¹². The impact origin for the Vredefort structure is still disputed¹⁷, with much debate centring on the unusual regional distribution of partly anomalous planar deformation features (PDFs¹⁸) in shocked quartz¹⁹. With the present poor knowledge of the shock-wave behaviour of minerals at a very high pre-shock temperature, the ongoing debate cannot be resolved. Largely for this reason, we have studied shock effects in heated quartz which is the most widely occurring and most intensely studied index mineral for shock-wave barometry.

Our shock-loading experiments were done on disks of gem-quality single-crystal quartz < 0.5 mm thick, in sample holders consisting of a cylindrical ARMCO-Fe container; this arrangement allows complete recovery of the sample. The system was externally heated in a muffle furnace, and immediately before the passage of the shock wave, the quartz samples were at 630 ± 5 °C. To generate the shock waves we used a high-explosive device similar to one described previously²⁰, but the flyer plate was thermally isolated by a combination of Al-, Kapton- and Mylar-foils. Peak shock pressures were 20–48 GPa (precision $\pm 2\%$) with the direction of shock-wave propagation perpendicular to the {10 $\bar{1}$ 0} face of quartz. The pre-shock temperature, which is well in the range given above for the Sudbury and Vredefort structures, clearly exceeds 573 °C, the α - β transition temperature of quartz. Consequently, the samples had the hexagonal β -quartz structure during shock compression.

The recovered quartz samples were investigated by spindle²¹ and universal stage techniques as well as by Guinier-Jagodzinski X-ray methods. Figure 1 shows the shock-induced variation of refractive indices in quartz, shocked at 630 °C and at room temperature using an identical experimental arrangement^{22,23}. Below 25 GPa, both types of shocked samples have similar optical behaviour, and the measured properties are very similar to those of the unshocked reference quartz. Yet in the small pressure range from 25 to 26 GPa, the refractive indices of heated quartz decrease markedly and discontinuously, whereas unheated quartz shows an equivalent change between 25 and 35 GPa (refs 7, 9, 22, 23). The complete transformation to diaplectic glass is achieved at 26 GPa for heated quartz, whereas quartz shocked at room temperature becomes completely isotropic only at pressures between 35 and 38 GPa (ref. 22).

The distinctly different optical characteristics of unheated and heated shocked quartz are also obvious from the orientation of PDFs shown in Fig. 2. At a shock pressure of 20 GPa, the

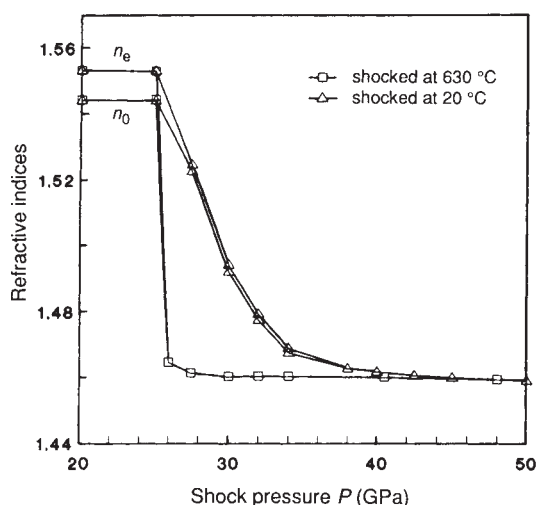


FIG. 1 Refractive indices n_e and n_o as a function of shock pressure for quartz experimentally shocked at 630 °C and at room temperature using a set-up with flyer plates driven by high explosives^{22,23}; measurements were made with a spindle stage.

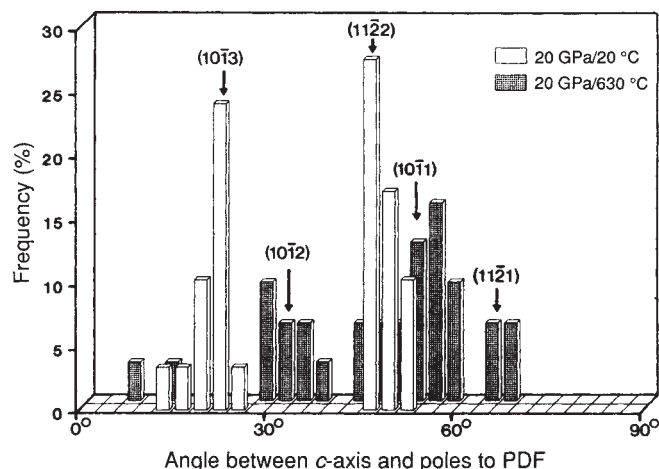


FIG. 2 Comparison of orientations of planar deformation features in quartz shocked experimentally at 630 °C and at room temperature; measurements were made with a universal stage. In the unheated quartz, poles to planar features are concentrated at {10 $\bar{1}$ 3}, and near {11 $\bar{2}$ 2} and {10 $\bar{1}$ 1}. Other poles are given for reference only. Estimated error is $\pm 5^\circ$.

frequency distribution of PDFs in unheated quartz has two distinct maxima. From previous investigations on experimentally⁷ and naturally^{24,25} shocked quartz, the maximum at $\{10\bar{1}3\}$ seems fairly typical for a shock pressure of ~ 20 GPa. The other two favoured orientations, $\{11\bar{2}2\}$ and $\{10\bar{1}1\}$ both at a similar angle to the c -axis (47.7° and 51.8°), are more frequent than in quartz shocked in nature and in other experiments. This apparent discrepancy is due to the limited number of PDFs found in the few quartz grains available for this investigation, which were shocked in a previous series of experiments²³. Although the low-temperature results may not be representative in detail of what happens in nature, this does not affect our main, qualitative conclusions on high-temperature shock metamorphism. In contrast, PDFs in heated quartz show a broad distribution of crystallographic orientations with indistinct maxima which are difficult to index. For PDFs of heated quartz, complete absence of the $\{10\bar{1}3\}$ -orientation is most characteristic, and orientations with high angles to the c -axis occur frequently. The difference in orientation of PDFs in unheated and heated quartz shocked to the same peak pressure may reflect the slight difference between the lattices of α and β -quartz. This conclusion may be supported by the fact that quartz grains from a quartzite shocked at 440°C are still dominated by PDFs parallel to $\{10\bar{1}3\}$ (ref. 6).

X-ray data confirm the optical observations. Diffraction patterns of three quartz samples shocked at 20°C and at 630°C were selected to demonstrate the transformation from crystalline quartz to X-ray amorphous diaplectic glass. The following conclusions can be drawn from Fig. 3: first, the X-ray diffraction patterns for unheated and heated quartz (both shocked at 25 GPa) are identical and display the complete set of reflections well known for unshocked quartz; second, unheated quartz

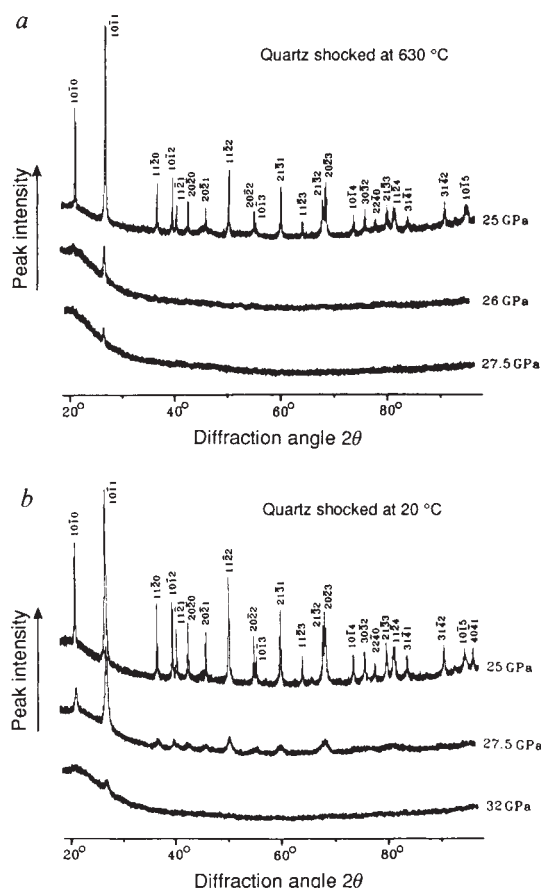


FIG. 3 Variation of X-ray diffraction patterns with shock pressure for quartz experimentally shocked (a) at 630°C and (b) at room temperature (Guinier-Jagodzinski method).

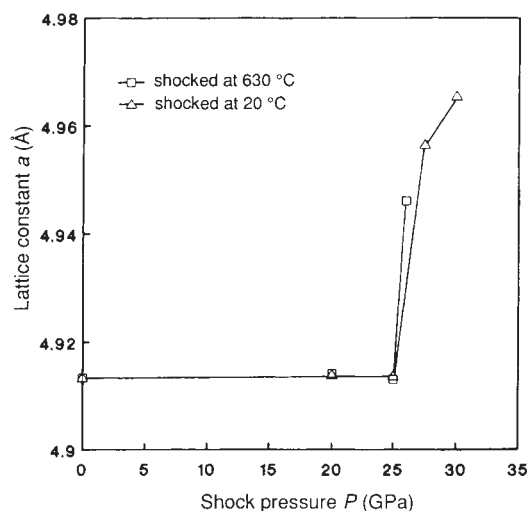


FIG. 4 Lattice constant a as a function of shock pressure in quartz experimentally shocked at 630°C and at room temperature.

shocked to 27.5 GPa still shows all front reflections, and only a few back reflections are absent, which can be explained by a loss of long-range order. The observed lowered peak intensities and the X-ray line broadening are well-known shock effects in quartz²⁶. The heated quartz shocked at 27.5 GPa, however, displays neither back reflections nor front reflections except for one very weak and broad line.

From the X-ray diffraction spectra, we determined d -spacings and lattice parameters. As a representative example for the shock-induced breakdown of the crystal structure, the lattice constant a is plotted as a function of shock pressure in Fig. 4. Below 25 GPa this parameter remains unchanged from that of the reference sample for both unheated and heated quartz. Above 26 GPa, the crystal structure of the heated quartz breaks down, and because of amorphization, lattice parameters are no longer measurable. In contrast, unheated quartz shows this effect only above 30 GPa. Moreover, above 25 GPa, heated quartz shows a stronger change in a with increasing pressure than does the unheated sample (Fig. 4). The same effect was found for the lattice parameter c . Obviously, the lattice expansion results from the pressure release after the shock wave has passed²⁷.

The strong dependence of shock effects in quartz on the pre-shock temperature, as indicated by our very consistent optical and X-ray results has far-reaching implications. Shock-wave barometers currently in use are based almost exclusively on data from recovery experiments done at room temperature. Our investigation indicates that the application of these data to natural impactites must be limited to low-temperature target lithologies. Moreover, our results for experimentally shocked heated quartz show that systematic recovery experiments on all major rock-forming minerals at variable ambient temperatures are essential if we are to understand the complete pressure-temperature regime of impact metamorphism of planetary bodies; the results may be unexpected. The characteristics of high-temperature shock metamorphism, still largely unknown, also bear on the recognition and understanding of impact phenomena resulting from basin-sized impacts into the early terrestrial crust. Further experiments will help to improve the present mechanical models for the formation of very large impact basins, which affect crustal sections tens of kilometres thick with vertical temperature gradients of several hundred degrees.

Received 28 November 1991; accepted 25 February 1992.

1. Roddy, D. J., Pepin, R. O. & Merrill, R. B. (eds) *Impact and Explosion Cratering* (Pergamon, New York, 1977).

2. Melosh, H. J. *Impact Cratering: A Geological Process* (Oxford University Press, New York, 1989).
3. Scott, E. R. D. et al. in *Asteroids II* (eds Binzel, R. P., Gehrels, T. & Matthews, M. S.) 701–739 (University of Arizona Press, Tucson, 1989).
4. Stöffler, D., Keil, K. & Scott, E. R. D. *Geochim. cosmochim. Acta* **55**, 3845–3867 (1991).
5. Reimold, W. U. *Lunar planet. Sci.* **XIX**, 970–971 (1990).
6. Huffman, A. R., Brown, J. M. & Carter, L. in *Shock Compression of Condensed Matter* (eds Schmidt, S. C., Johnson, J. N. & Davison, L. W.) 649–652 (Elsevier, Amsterdam, 1990).
7. Hörz, F. in *Metamorphism of Natural Materials* (eds French, B. M. & Short, N. M.) 243–253 (Mono, Baltimore, 1968).
8. Stöffler, D. *Fortschr. Miner.* **49**, 50–113 (1972).
9. Stöffler, D. *J. non-cryst. Solids* **67**, 465–502 (1984).
10. Reimold, W. U. & Hörz, F. *Lunar planet. Sci.* **XVII**, 703–704 (1986).
11. Stöffler, D., Bischoff, A., Buchwald, V. & Rubin, A. E. in *Meteorites and the Early Solar System* (eds Kerridge, J. F. & Matthews, M. S.) 165–202 (University of Arizona Press, Tucson, 1988).
12. Schreyer, W. *J. Petrol.* **24**, 26–41 (1983).
13. Fricke, A., Medenbach, O. & Schreyer, W. *Tectonophysics* **171**, 169–183 (1990).
14. Faggart, B. E., Basu, A. R. & Tatsumoto, M. *Science* **230**, 436–439 (1985).
15. Lakomy, R. *Meteoritics* **25**, 195 (1990).
16. Grieve, R. A. F., Stöffler, D. & Deutsch, A. *J. geophys. Res.* **96**, 22753–22764 (1991).
17. Nicholaysen, L. O. & Reimold, W. U. (eds) *Tectonophysics* **171** (spec. issue) (1990).
18. Grieve, R. A. F., Sharpton, V. L. & Stöffler, D. *Eos* **71**, 1792–1793 (1990).
19. Grieve, R. A. F., Coderre, J. M., Robertson, P. B. & Alexopoulos, J. *Tectonophysics* **171**, 185–200 (1990).
20. Müller, W. F. & Hornemann, U. *Earth planet. Sci. Lett.* **7**, 251–264 (1969).
21. Medenbach, O. *Fortschr. Miner.* **63**, 111–133 (1985).
22. Rehfeldt-Oskierski, A. & Stöffler, D. *Lunar planet. Sci.* **XVII**, 697–698 (1986).
23. Grothues, J., Deutsch, A., Hornemann, U. & Stöffler, D. *Lunar planet. Sci.* **XX**, 365–366 (1989).
24. Von Engelhardt, W. & Bertsch, W. *Contrib. Miner. Petrol.* **20**, 203–234 (1969).
25. Robertson, P. B. & Grieve, R. A. F. in *Impact and Explosion Cratering* (eds Roddy, D. J., Pepin, R. O. & Merrill, R. B.) 687–702 (Pergamon, New York, 1977).
26. Hörz, F. & Quaide, W. L. *The Moon* **6**, 45–82 (1972).
27. Ashworth, J. R. & Schneider, H. *Phys. Chem. Min.* **11**, 241–249 (1985).

ACKNOWLEDGEMENTS. We thank U. Franz, H. P. Bernauer, and B. Nasdala for technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft, Bonn.

Alleghenian age of the Upper Mississippi Valley zinc–lead deposit determined by Rb–Sr dating of sphalerite

Joyce C. Brannon*, Frank A. Podosek* & Roger K. McLimans†

* Department of Earth and Planetary Sciences, and McDonnell Center for the Space Sciences, Washington University, St Louis, Missouri 63130–4899, USA

† E. I. du Pont de Nemours and Company, Incorporated, Wilmington, Delaware 19898, USA

MISSISSIPPI Valley type (MVT) ore deposits in North America are generally thought to have formed from warm brines which migrated from sedimentary basins to sites of deposition in sedimentary strata of the neighbouring craton long after lithification of the sediments^{1–3}. When and how these migrations occurred is not well understood, but recent geological evidence suggests that continental-scale brine migration can be triggered by tectonic collisions^{4–6}. Indirect evidence^{7–9} has suggested an association of North American MVT mineralization with the Late Palaeozoic Alleghenian/Ouachita orogeny; recently, however, Nakai *et al.*¹⁰ obtained a Rb–Sr age of 377 ± 29 Myr for sphalerites from the East Tennessee MVT district, which they linked to the Middle Palaeozoic Acadian orogeny. Here we report Rb–Sr dates for sphalerites from the MVT 'type section' — the Upper Mississippi Valley (UMV) zinc–lead district. Rb–Sr data from two distinct sphalerite bands from the West Hayden ore body yield two independent but indistinguishable isochrons of Permian age (269 ± 6 Myr and 270 ± 4 Myr), consistent with a brine migration episode initiated by the Alleghenian/Ouachita orogeny. In conjunction with the previous data indicating a link to the Acadian orogeny^{10,11}, our results indicate that North American MVT deposits formed in at least two distinct tectonic episodes.

As any genetic model must involve an appropriate temporal relation between cause and effect, the hypothesis that MVT deposits result from large-scale migration of basinal brines initiated by tectonic orogeny can be tested by comparing the time

of ore formation with the time of the associated orogeny. Although the importance of such chronological constraints is recognized and has inspired considerable experimental work, the results available at present are largely inconclusive, as the available chronological data are generally difficult to relate to ore formation.

As an example, palaeomagnetic studies on Palaeozoic strata in the mid-continent of North America indicate, almost without exception, a resetting of palaeopole position during Pennsylvanian–Permian time (refs 6–8, and as reviewed by Elmore and McCabe⁹). Because the palaeomagnetic signal can be set or reset by fluid-rock interactions, resulting in precipitation or dissolution and reprecipitation of magnetic carriers such as authigenic framboidal magnetite^{9,12,13}, these palaeomagnetic ages are commonly believed to be related to MVT ore deposits. If this inference is correct it supports an association^{7–9} of mid-continent MVT ore deposits with the Alleghenian/Ouachita orogeny, which occurred during Pennsylvanian–Permian time. Historically, the ubiquity of Pennsylvanian–Permian ages in the palaeomagnetic data has fostered the view that essentially all mid-continent North American MVT deposits were formed at the same time.

Radiometric ages have also been obtained on authigenic minerals^{6,14–18} (such as glauconite, K-feldspar, illite and sericite) in mid-continent Palaeozoic strata, which are also inferred to be formed in fluid-rock interactions and are thus considered to be associated with MVT ore deposits. Such ages tend to fall into three groups¹⁴: Permian, and so consistent with the palaeomagnetic ages and an Alleghenian/Ouachita association, but also late Devonian and early Devonian, times possibly correlating with earlier North American orogenies¹⁵. Often, different ages are obtained at the same geographical site, sometimes for the same mineral.

Evidently, the palaeomagnetic and the various isotopic systems are recording or being overprinted by different geological events, and even if these events are mediated by fluid-rock interactions it is unclear which, if any, of such fluids are the fluids responsible for the ore minerals. There is a well recognized association, based on fluid inclusion studies, between MVT ores and hydrothermal dolomites deposited throughout the region^{19,20}, and this association supports the hypothesis of long-distance fluid migration. These regionally distributed minerals, plausibly associated with MVT ores by similarity of fluid inclusion compositions, are as yet undated, and the palaeomagnetically and radiometrically dated minerals referred to above have not yet been unambiguously linked to MVT mineralization.

The direct approach of seeking ages for the main ore minerals themselves is obvious, but reliable data are scarce. A principal reason for the dearth of radiometric ages is technical: for the trace element systems used in isotopic geochronology, either the parent/daughter ratio is so low in these minerals that the system is inapplicable or the absolute abundances of the relevant elements are too low for reliable analysis. A possible additional problem, not yet well explored, is that variations in the isotopic composition of the daughter element in the ore-forming fluids may obscure the *in situ* accumulation of the radiogenic daughter isotope even when the isotopic analysis is technically feasible.

Recent improvements in analytical procedures have allowed at least the Rb–Sr isotope system to be applied to the main MVT ore minerals. Brannon *et al.*²¹ found large variations in ⁸⁷Sr/⁸⁶Sr among several minerals in the Viburnum Trend (southeastern Missouri), but concluded that variation in fluid composition was so great as to preclude chronological inferences. In contrast, Nakai *et al.*¹⁰ inferred a Rb–Sr isochron of Acadian age (377 ± 29 Myr ago) for sphalerite from the Coy Mine (eastern Tennessee). Nakai *et al.*¹¹ have also inferred Acadian ages for other widely separated North American MVT deposits. Such results suggest a prevalent Acadian association for North American MVT mineralization, in sharp contrast to the Alleghenian/Ouachita association more prevalent in other kinds

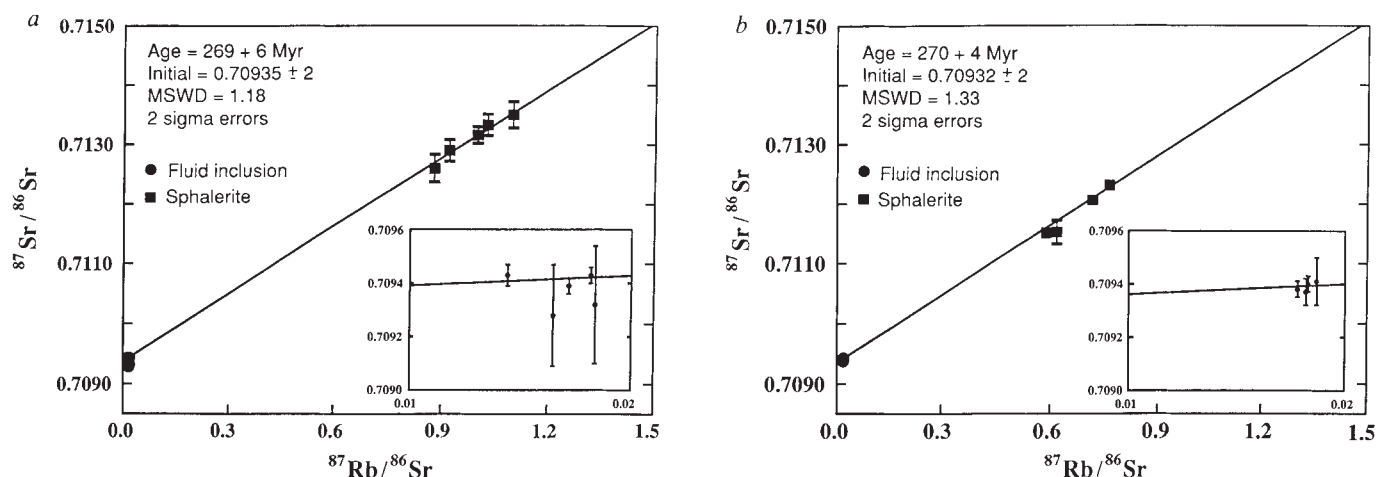


FIG. 1 Rb-Sr isochron diagrams for (a) Zone B horizon 17 and (b) Zone C horizon 23 in the West Hayden ore body in the Upper Mississippi Valley zinc-lead district. Data from Table 1. Insets show low Rb/Sr fluid inclusion data. Isochron fits calculated by the York-fit Model 1 algorithm of Ludwig³⁰. Ages calculated for $\lambda = 1.42 \times 10^{-11} \text{ yr}^{-1}$. A fit to only the 'solids' data

(excluding the 'fluid inclusions' data) yields an age of $284 \pm 20 \text{ Myr}$ and initial $^{87}\text{Sr}/^{86}\text{Sr} = 0.70916 \pm 22$ with MSWD (mean standard weighted deviation) = 1.0. A fit to all the data from both horizons yields an age of $269 \pm 4 \text{ Myr}$ and initial $^{87}\text{Sr}/^{86}\text{Sr} = 0.70933 \pm 2$ with MSWD = 1.4.

of chronological data. Here we report chronologically relevant Rb-Sr data for the MVT 'type-section', the Upper Mississippi Valley (UMV) zinc-lead district.

The UMV district is located on an uplifted area on the southwest flank of the Wisconsin arch; it is surrounded by the Wisconsin dome to the north, the Forrest City basin to the west, the Illinois basin to the south and the Michigan basin to the east. Regional tectonic deformation, almost all post-Silurian, occurred before ore deposition²²⁻²⁴. Minor faults of reverse, bedding-plane, normal and shear types are common, and a well-developed vertical and inclined joint system prevails throughout the area; the joint and fracture systems provide easy fluid flow paths between clastic aquifers and carbonate strata. Mineralization, which includes zinc, lead and iron sulphides, barite, dolomite and calcite, occurs dominantly in Middle

Ordovician strata. The most common ore bodies are pitch-and-flat structures and gash-vein deposits²³. Ore precipitation occurred principally as open-space filling with mineral growth progressing outward from the wall rock, such that there is a well-defined space-time relationship²⁵. The most abundant mineral deposited is sphalerite, which occurs in distinct bands of different colour. McLimans *et al.*²⁵ developed and correlated a sphalerite stratigraphy across the ore district, delineating three zones A, B and C (early to late, respectively) and several key marker horizons or bands within each zone.

We analysed samples from the West Hayden ore body, five cut as wafers from zone B horizon 17 and four cut from zone C horizon 23. Each sample was analysed in two fractions, a 'fluid inclusion' fraction obtained by washing in water after crushing to open fluid inclusions, and the complementary 'solid'

TABLE 1 Isotopic and elemental analyses of Upper Mississippi Valley sphalerites

Sample	Sample weight* (mg)	[K] p.p.m.†	[Ca] p.p.m.†	K/Ca (weight)	Ca/Sr (weight)	[Rb] p.p.b.†	[Sr] p.p.b.†	$^{87}\text{Rb}/^{86}\text{Sr} \ddagger$	$^{87}\text{Sr}/^{86}\text{Sr} \S \ddagger$
Sphalerite fluid inclusions (crushed sample washes)									
58B 17 1	69	0.8	6.3	0.12	31	1.3	203	0.0184 ± 2	0.70932 ± 22
17 2	213	1.2	10.3	0.12	31	2.1	330	0.0182 ± 2	0.70943 ± 3
17 3	154	0.7	7.5	0.10	30	1.2	248	0.0144 ± 2	0.70943 ± 4
17 4	100	0.7	6.7	0.11	30	1.2	208	0.0165 ± 2	0.70928 ± 19
17 5	270	0.7	6.6	0.11	32	1.2	207	0.0172 ± 2	0.70939 ± 3
10C 23 1	176	0.4	3.1	0.12	30	0.7	102	0.0187 ± 2	0.70943 ± 11
23 2	486	0.4	3.6	0.12	27	0.8	131	0.0178 ± 2	0.70938 ± 3
23 3	703	0.3	3.0	0.10	30	0.6	99	0.0183 ± 2	0.70940 ± 3
23 4	1,066	0.4	2.9	0.13	28	0.7	105	0.0182 ± 2	0.70937 ± 5
Sphalerite solids (wash residues)									
58B 17 1	69	1.7	3.2	0.54	68	18	47	1.099 ± 9	0.71350 ± 22
17 2	206	2.3	2.9	0.78	38	23	76	0.879 ± 6	0.71260 ± 23
17 3	154	2.1	2.8	0.74	41	24	69	1.000 ± 8	0.71316 ± 14
17 4	101	2.3	4.1	0.55	59	22	70	0.919 ± 6	0.71290 ± 18
17 5	270	2.0	2.7	0.73	44	22	62	1.026 ± 9	0.71333 ± 18
10C 23 1	201	1.1	2.4	0.50	77	7	31	0.618 ± 4	0.71152 ± 20
23 2	518	2.6	4.5	0.58	113	11	40	0.767 ± 5	0.71231 ± 6
23 3	740	1.4	2.3	0.60	66	9	35	0.719 ± 5	0.71206 ± 4
23 4	1,084	1.5	3.0	0.51	68	9	44	0.589 ± 4	0.71151 ± 11

* Aliquot used for Rb-Sr analysis.

† Concentrations by isotope dilution, based on total starting sample weight; p.p.b., parts per 10^9 .

‡ Two-standard-deviation errors.

§ Adjusted to 0.71014 for NBS-987.

fraction obtained by dissolution after removal of the fluid inclusion fraction. Details of the sample processing and trace element analytical protocols have been reported previously²¹; results for Rb, Sr, K and Ca are presented in Table 1.

Elemental ratios in the fluid inclusions are fairly similar, suggesting that the ore-forming fluids were well mixed. $^{87}\text{Rb}/^{86}\text{Sr}$ is sufficiently low in the fluid inclusions that *in situ* growth of ^{87}Sr is minor, and accurate identification of initial $^{87}\text{Sr}/^{86}\text{Sr}$ composition is possible for any plausible age. For each sample this ratio is higher than that in any Phanerozoic sea water and so indicates, as in other MVT deposits^{10,11,21} that the UMV ore-forming fluids must have interacted with the Precambrian basement or clastics derived from the basement. For each horizon, initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in the various fluid inclusion samples are indistinguishable. Initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in the two horizons differ by an amount close to the limit resolvable within analytical uncertainties and so could be either equal or slightly disparate, corresponding to slight evolution of the fluids between the two horizons. The (near-) uniformity of initial $^{87}\text{Sr}/^{86}\text{Sr}$ parallels the observed uniformity of sulphur isotopic composition²⁶ at UMV and is quite different from the strong variations observed in the Viburnum Trend²¹.

The present $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in the solid fractions are higher and more variable than those in the fluid inclusion fractions, and their correlation with Rb/Sr ratios is consistent with an isochron relationship (Fig. 1). Independent isochron fits to the data for each horizon separately yield indistinguishable ages, 269 ± 6 Myr for horizon 17 and 270 ± 4 Myr for horizon 23. Indistinguishable ages also result from fits to all the data from both horizons considered together, and to the solids data alone without the fluid inclusion data (see Fig. 1). Mixing relationships involving K or Ca with Rb and Sr are not linear, supporting the interpretation that the relationships evident in Fig. 1 are isochrons and not merely mixing lines. We thus infer that these data determine the time of formation of these UMV horizons.

Doe *et al.*²⁷ studied the U-Th-Pb system in Precambrian basement granites recovered from the UPH drill hole located at the southeastern edge of the UMV district. They found evidence for substantial lead loss to what they suggested were plausible MVT ore-forming fluids at a 'most probable age' of 260 ± 35 Myr (lower concordia intercept). Thus both the strontium (this work) and lead²⁷ isotope systems indicate interaction of basinal brines with the basement and both systems suggest compatible times. Pennsylvanian-Permian palaeomagnetic ages are consistent with these isotopic ages. In addition, Bethke²⁸ has advanced a model for gravity-driven groundwater flow from the Illinois basin thought to have been initiated by the early Permian uplift of the Pascola Arch in response to the Alleghenian/Ouachita orogeny^{5,29}. Taken together, these considerations provide a strong case for Illinois basin fluid migration leading to UMV ore deposition ~ 270 Myr ago.

These results contrast sharply with the Acadian ages found by Nakai *et al.*^{10,11} but they do not conflict with them because they refer to different locations. Taken together, these results indicate that North American MVT deposits formed in multiple (at least two) distinct episodes, plausibly linked to major tectonic events. □

15. Duffin, M. E., Lee, M., de V. Klein, G. & Hag, R. L. *J. sedim. Petrol.* **59**, 848-861 (1989).
16. Posey, H. H., Stein, H. J., Fullagan, P. D. & Kish, S. A. in *Int. Conf. MVT-type Pb-Zn Deposits* (eds Kisvarsanyi, G. *et al.*) 166-173 (University of Missouri, 1983).
17. Laskowski, T. E., Fluegeman, R. H. & Grant, N. K. *Geology* **8**, 368-370 (1980).
18. Marshall, B. D., Woodard, H. H. & DePaolo, D. J. *Geology* **14**, 936-938 (1986).
19. Coveney, R. M. Jr., Goebel, E. D. & Ragan, V. M. *Econ. Geol.* **82**, 740-751 (1987).
20. Leach, D. L. *Econ. Geol.* **74**, 931-937 (1979).
21. Brannon, J. C. *et al. Geochim. cosmochim. Acta* **55**, 1407-1419 (1991).
22. Heyl, A. V., Agnew, A. F., Lyons, E. J. & Behre, C. H. *J. U.S. geol. Surv. Prof. Pap.* 309 (1959).
23. Heyl, A. V. in *Ore Deposits of the United States 1933-1967* (ed. Ridge, J. D.) 431-459 (Mining Metal. Petrol. Engineers, New York, 1968).
24. Heyl, A. V. in *Int. Conf. MVT-type Pb-Zn Deposits* (eds Kisvarsanyi, G. *et al.*) 27-60 (University of Missouri, 1983).
25. McLimans, R. K., Barnes, H. L. & Ohmoto, H. *Econ. Geol.* **75**, 351-361 (1980).
26. McLimans, R. K. thesis, Pennsylvania State Univ. (1977).
27. Doe, B. R., Stuckless, J. S. & Delevaux, M. H. *J. geophys. Res.* **88**, 7335-7345 (1983).
28. Bethke, C. M. *Econ. Geol.* **81**, 233-249 (1986).
29. Viele, G. W. *Geol. Soc. Am. Abstr. Prog.* **15**, 712 (1983).
30. Ludwig, K. R. *U. S. geol. Surv. Open-file Rep.* 88-557 (1988).

Evidence from magnetic fabric for the flow pattern of magma in the Mackenzie giant radiating dyke swarm

Richard E. Ernst* & W. R. A. Baragar†

* Ottawa-Carleton Geoscience Centre, Department of Geology, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada

† Geological Survey of Canada, 601 Booth Street, Ottawa, Ontario K1A 0E8, Canada

CONTINENTAL flood basalts, and their associated swarms of mafic dykes, are thought to result from melting in mantle plumes¹⁻³, but the processes by which the resulting magma makes its way from mantle to crust remain obscure. Here we use magnetic anisotropy measurements to reconstruct the pattern of magma flow in the Proterozoic Mackenzie dyke swarm. Our results indicate that magma was injected vertically within 500 km of the focal point of the dyke swarm, and then travelled horizontally at all further distances out to at least 2,100 km. The sharp transition from vertical to horizontal flow at between 500 and 600 km may correspond to the outer boundary of melt generation in the mantle plume believed to be responsible for the Mackenzie igneous events.

Dykes are the primary conduit for mafic magma transport into the crust from a source area in the mantle. Although it is obvious that the overall direction of magma movement is upwards, the actual path may include lateral components of movement as well. In particular, studies of small dyke swarms by seismology, texture, geochemistry and magnetic susceptibility reveal variable but mainly lateral flow away from coeval volcanic or plutonic centres⁴⁻⁷. Mega-swarms (extending hundreds of kilometres along strike) found in shield areas are observed in a number of cases to fan away from coeval volcanic or plutonic centres, a prominent example being the Middle Proterozoic Mackenzie swarm of the Canadian Shield⁸. The fan-like geometry, along with geochemical evidence, suggests horizontal injection in dykes of this scale as well⁹⁻¹².

The case for lateral flow in mega-swarms can be rigorously tested by direct measurements of the flow fabric. In mafic dykes, however, the petrographic fabric may be only weakly present and difficult to map; indeed, the Mackenzie swarm is non-porphyrific and lacks other obvious flow indicators. Fortunately, subtle fabrics may be revealed by magnetic fabric analysis¹³⁻¹⁵. The magnetic fabric (anisotropy of magnetic susceptibility, AMS) of mafic magmas has been shown¹⁶ to reflect the rock fabric (particularly tabular or bladed plagioclase) because the distribution of late-crystallizing magnetite is determined by the pre-existing silicate fabric of the rock, and susceptibility anisotropy is a function of the anisotropic distribution of

Received 18 November 1991; accepted 25 February 1992.

1. White, D. E. *Econ. Geol.* **63**, 301-335 (1968).
2. Ohle, E. L. *Econ. Geol.* **75**, 161-172 (1980).
3. Sverjensky, D. A. *A. Rev. Earth planet. Sci.* **14**, 177-199 (1986).
4. Oliver, J. *Geology* **14**, 99-102 (1986).
5. Leach, D. L. & Rowan, E. L. *Geology* **14**, 931-935 (1986).
6. Bethke, C. M. & Marshak, S. A. *Rev. Earth planet. Sci.* **18**, 287-315 (1990).
7. Beales, F. W., Carracedo, J. C. & Strangway, D. W. *Can. J. Earth Sci.* **11**, 211-223 (1974).
8. Symons, T. A. & Sangster, D. F. *Econ. Geol.* **86**, 1-12 (1991).
9. Elmore, R. D. & McCabe, C. *Rev. Geophys.* **29**, 377-383 (1991).
10. Nakai, S., Halliday, A. N., Kesler, S. E. & Jones, H. D. *Nature* **346**, 354-357 (1990).
11. Nakai, S. *et al. Geochim. cosmochim. Acta* (submitted).
12. Lu, G., Marshak, S. & Kent, D. V. *Earth planet. Sci. Lett.* **99**, 351-361 (1990).
13. Jackson, M., McCabe, C., Ballard, M. M. & Vander Voo, R. *Geology* **16**, 592-595 (1988).
14. Hay, R. L., Lee, M., Kolata, D. R., Matthews, J. C. & Morton, J. P. *Geology* **16**, 743-747 (1988).

magnetite. AMS analysis is sensitive to fabrics with an anisotropy of only a few per cent. Mafic magma flowing through a dyke will acquire a preferential magnetic fabric with the minimum axis aligned perpendicular to the wall and a maximum axis in the direction of flow⁷.

The huge, fan-shaped, Mackenzie mafic dyke swarm, distributed over more than half of the Canadian Shield, converges toward a region of coeval volcanic and plutonic activity represented by the Coppermine River and Ekalulia basalts, and the Muskox layered intrusion (Fig. 1). Precise U-Pb baddeleyite ages indicate that the Muskox intrusion, several main sills and the Mackenzie dyke swarm were emplaced within a period of less than 5 million years beginning at 1,272 Myr ago, with most dykes injected at $1,267 \pm 2$ Myr (refs 17, 18).

Here we use AMS measurements to interpret the flow pattern in the Mackenzie swarm. Magnetic fabric measurements were made on 568 samples (at least two specimens each) from 51 sites representing 44 dykes collected at the following distances from the focus of the Mackenzie swarm: 400, 500, 600, 800, 1,000, 1,400 and 2,100 km (Fig. 1). The southern edge of the preserved remnants of the Coppermine River flood basalts is ~400 km from the swarm focus. Dyke thickness ranges from 10 to 90 m, and most samples were collected within a few metres of the dyke margins where the flow velocity gradient would have been largest and the chances of obtaining a preserved flow fabric the greatest. All samples were oriented with a solar compass and their anisotropy measured on a Sapphire Instruments SI-1. Samples from 400 and 500 km were also measured on a Kappa-bridge KLY-1 instrument, and identical results were obtained.

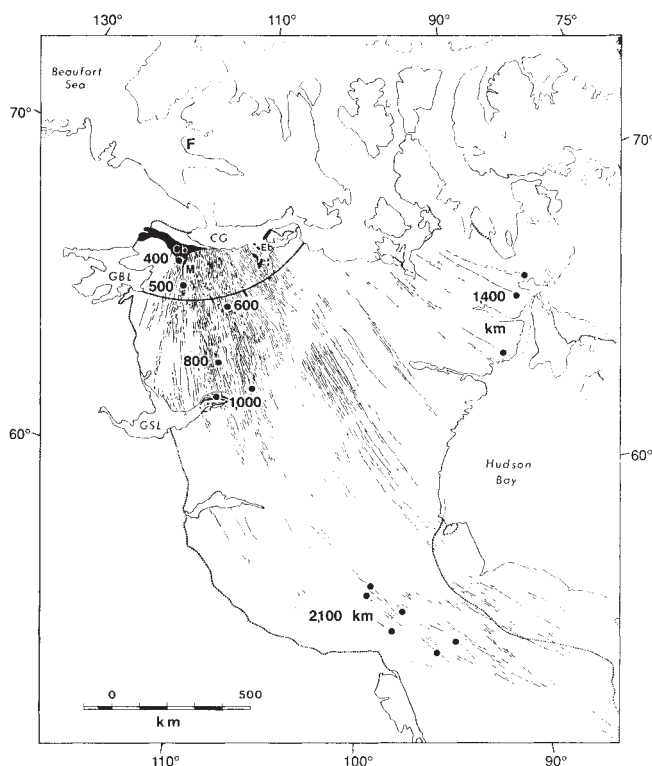


FIG. 1 Map of sample locations showing regional distribution of the Mackenzie dyke swarm (after Fahrig⁵) and coeval igneous rocks in the northwestern Canadian Shield (after LeCheminant and Heaman¹⁷). F is the focal point of the Mackenzie swarm; M, Muskox intrusion; Cb, Coppermine River basalts; Eb, Ekalulia basalts; CG, Coronation Gulf; GBL, Great Bear Lake; GSL, Great Slave Lake: Circular arc between 500 and 600 km marks the proposed outer boundary of melt generation in the mantle plume as interpreted from our AMS data. Heavy, ornamented line marks the boundary of Phanerozoic cover-rocks on the mainland. The location of the focus, F, was determined from converging great-circle dyke trends and is located with an uncertainty of less than 50 km.

Average magnetic anisotropy was 3%, and results from different specimens taken from the same sample generally agreed to within a few degrees. Near-surface specimens (from field-drilled cores) often had an anomalous fabric, probably because of surface weathering, and we therefore used only the data of specimens from depths greater than a few centimetres (usually >5 cm) for interpretation. The contoured data are shown in Figs 2 and 3. To eliminate the effects of local differences in dyke orientation, all AMS data were recalculated to a common dyke plane (arbitrarily chosen as north-south and vertical). The fabric

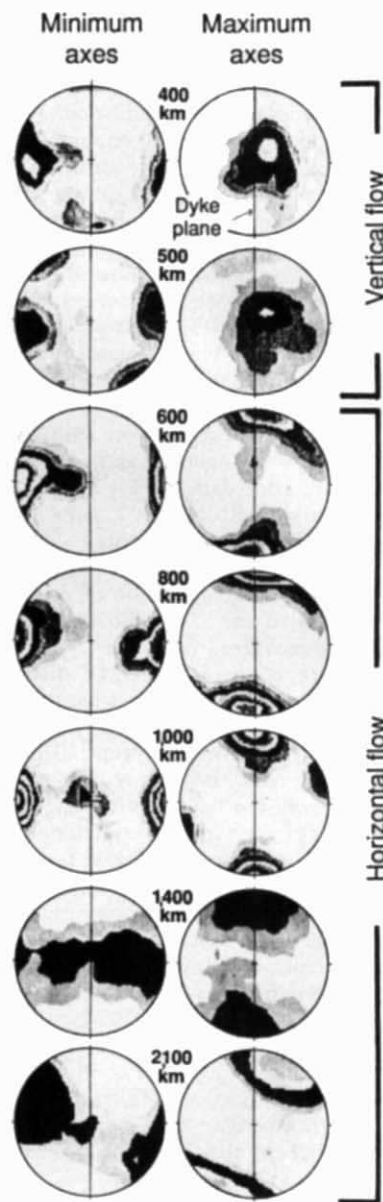


FIG. 2 Anisotropy of magnetic susceptibility (AMS) data from each sampling region contoured using the Kamb contouring method²⁶. Shown are equal-area projections for the minimum and maximum axes of the triaxial ellipsoid of magnetic susceptibility magnitude. Two specimens represent each sample. The number of samples, sites and dykes from each region are as follows: 400 km (77 samples, 8 sites, 6 dykes), 500 km (80 samples, 6 sites, 4 dykes), 600 km (100 samples, 9 sites, 9 dykes), 800 km (78 samples, 6 sites, 6 dykes), 1,000 km (98 samples, 10 sites, 9 dykes), 1,400 km (39 samples, 5 sites, 4 dykes), 2,100 km (15 samples from 1 site, 1 dyke). See Fig. 3 for data from 5 other dykes from the 2,100 km distance. The samples from 1,400 km have been described by Park²⁷. AMS results from 400, 500 and 1,400 km were published in ref. 28.

in Figs 2 and 3 is considered primary because the mineralogy and, in particular, the opaque oxides have been generally only weakly affected by deuteric alteration, and palaeomagnetic studies of the swarm^{19,20} have consistently yielded stable primary remanence directions.

For most samples at distances of 400, 500, 600, 800 and 1,000 km, for many samples at 1,400 km and for one dyke at 2,100 km, the minimum susceptibility axes cluster in a direction perpendicular to the dyke plane (Fig. 2). This is a pattern consistent with a magma flow origin, and the inclination of the maximum axis is interpreted to equal the inclination of magma flow. We investigated the possibility that the polarity of flow could be determined from imbrication of the fabric near the dyke walls⁷; initial results are inconclusive.

At the distances of 400 and 500 km, the magnetic fabric data indicate dominantly vertical magma flow, whereas data from distances of 600, 800, 1,000, 1,400 and 2,100 km indicate horizontal flow (Fig. 2). At 1,400 km, the minimum axes are poorly grouped and only weakly aligned perpendicular to the dyke plane; this may be due partly to unknown dyke-dip corrections, but may also have an independent cause seen more clearly in the data from 2,100 km. One dyke from the 2,100-km locality (Fig. 2) has the pattern indicative of horizontal flow whereas the five other dykes (Fig. 3) have an anomalous magnetic fabric characterized by vertically oriented minimum axes. Park and others²¹ recorded vertical minimum axes in the Proterozoic Mealy dyke swarm of Labrador, with maximum axes showing a preferred orientation parallel to the strike of the dyke swarm. They attributed the fabric to vertical compaction in a static magma column with the direction of minimum stress along the dyke direction. Fabric orientation of the 2,100-km dykes (Fig. 3) might be similarly explained, but lack of a preferred direction for the maximum axes would indicate the presence of a quasi-equal horizontal stress field at the time of cooling. This type of fabric is best developed in the most distal Mackenzie sites at 2,100 and 1,400 km, perhaps implying that at this distance (close to the maximum extent of the swarm) low-energy flow in the dykes produced only a weak fabric that was easily overprinted by the effects of post-flow compaction (possibly by thermal contraction, crystal settling, filter pressing or plastic deformation).

The Mackenzie dykes thus provide direct evidence for long-distance lateral magma flow in a mega-swarm; lateral flow began between 500 and 600 km from the apex of the swarm and continued outwards to a distance of >2,100 km. Lateral flow on this scale is consistent with recent fluid mechanical/thermal theory and modelling^{22,23}.

If the interpretation of predominantly horizontal flow in major dyke swarms is correct, then one implication is that dyke magmas are derived from mantle underlying only the near-focal region of a swarm (that which shows vertical flow) and not its entire extent. So despite the prolific dissection of shield areas by mega-dyke swarms, it must be borne in mind that they probably represent only scattered and local sampling of the underlying mantle.

The transition from vertical to horizontal flow occurs in the interval between 500 and 600 km from the focus (F) of the Mackenzie swarm. It is tempting to speculate that this transition marks the outer boundary of the mantle plume proposed as the source of the Mackenzie igneous events¹⁷. If this were so, however, the implied plume diameter of ~1,000 km would be half or less of that predicted by Griffiths and Campbell from their experiments on plume analogues^{1,2,24}. These indicate that plumes originating at the mantle-core boundary achieve diameters of 1,000 km in free ascent, but flatten to disks of over 2,000 km diameter when they impinge on the lithosphere. Probably, therefore, the zone of vertical flow in the focal region of the Mackenzie dykes marks only the region of magma generation within a plume, which is smaller than the plume itself. In certain of their experiments, Griffiths and Campbell^{2,24} observed a zone

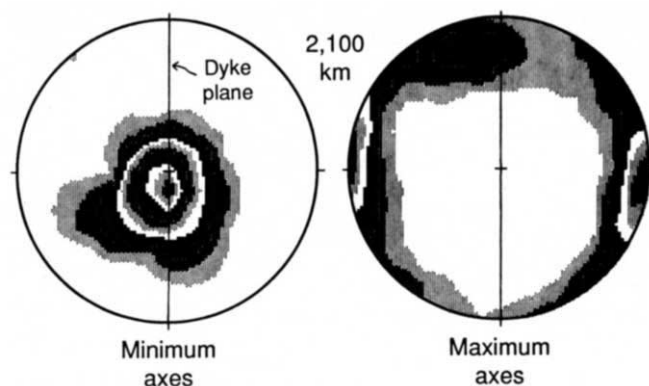


FIG. 3 Overprint fabric at 2,100 km distance. Data from 81 samples, 6 sites, 5 dykes. Stereonets as in Fig. 2.

of instability within a region containing the flattened plume head and overlying boundary zone, and this suggested to them a mechanism by which flood basalts could be generated. Flood basalts are thought to derive from depths of 40–60 km (ref. 25), which is closer to the surface than plume heads can rise without extraordinary thinning of the lithosphere, such as might result from the coincidence of a spreading ridge and plume centre³. But instability within the plume head may enable the plume to rise into the melt zone by displacing part of the overlying mantle (the 'squeeze zone') and to generate there, by partial melting, the voluminous magmas required. Under these circumstances, the melt zone may occupy only a portion of the plume head, which could be analogous to the zone of vertical flow identified here in the focal region of the Mackenzie dykes. □

Received 19 December 1991; accepted 10 March 1992.

- Griffiths, R. W. & Campbell, I. H. *Earth planet. Sci. Lett.* **99**, 66–78 (1990).
- Campbell, I. H. & Griffiths, R. W. *Earth planet. Sci. Lett.* **99**, 79–93 (1990).
- White, R. & McKenzie, D. *J. geophys. Res.* **94**, 7685–7729 (1989).
- Baer, G. & Reches, Z. *Geology* **15**, 569–572 (1987).
- Sigurdsson, H. in *Mafic Dyke Swarms* (eds Halls, H. C. & Fahrig, W. F.) 55–64 *Geol. Assoc. Canada Spec. Pap.* 34 (1987).
- Smith, R. P. in *Mafic Dyke Swarms* (eds Halls, H. C. & Fahrig, W. F.) 47–54 *Geol. Assoc. Canada Spec. Pap.* 34 (1987).
- Knight, M. D. & Walker, G. P. L. *J. geophys. Res.* **93**, 4301–4319 (1988).
- Fahrig, W. F. in *Mafic Dyke Swarms* (eds Halls, H. C. & Fahrig, W. F.) 331–348 *Geol. Assoc. Canada Spec. Pap.* 34 (1987).
- Gibson, I. L., Sinha, M. N. & Fahrig, W. F. in *Mafic Dyke Swarms* (eds Halls, H. C. & Fahrig, W. F.) 109–121 *Geol. Assoc. Canada Spec. Pap.* 34 (1987).
- Halls, H. C. in *Mafic Dyke Swarms* (eds Halls, H. C. & Fahrig, W. F.) 483–492 *Geol. Assoc. Canada Spec. Pap.* 34 (1987).
- MacDonald, R., Wilson, L., Thorpe, R. S. & Martin, A. *J. Petrol.* **29**, 559–583 (1988).
- Ernst, R. E. & Bell, K. *J. Petrol.* (in the press).
- Hrouda, F. *Geophys. Surv.* **5**, 37–82 (1982).
- MacDonald, W. D. & Ellwood, B. B. *Rev. Geophys.* **25**, 905–909 (1987).
- Potter, D. K. & Stephenson, A. *Geophys. Res. Lett.* **15**, 1097–1100 (1988).
- Hargraves, R. B., Johnson, D. & Chan, C. Y. *Geophys. Res. Lett.* **18**, 2193–2196 (1991).
- LeCheminant, A. N. & Heaman, L. M. *Earth planet. Sci. Lett.* **96**, 38–48 (1989).
- LeCheminant, A. N. & Heaman, L. M. *Geol.-Mineral. Assoc. Canada Ann. Meeting Prog. Abstr.* **16**, A73 (1991).
- Fahrig, W. F. & Jones, D. L. *Can. J. Earth Sci.* **6**, 679–688 (1969).
- Buchan, K. L. & Halls, H. C. in *Mafic Dykes and Emplacement Mechanisms* (eds Parker, A. J., Rickwood, P. C. & Tucker, D. H.) 209–230 (Balkema, Rotterdam, 1990).
- Park, J. K., Tanczyk, E. I. & Desbarats, A. *J. geophys. Res.* **93**, 13689–13704 (1988).
- Bruce, P. M. & Huppert, H. E. *Nature* **342**, 665–667 (1989).
- Lister, J. R. & Kerr, R. C. *J. geophys. Res.* **96**, 10049–10077 (1991).
- Griffiths, R. W. & Campbell, I. H. *J. geophys. Res.* **96**, 18295–18310 (1990).
- Jaques, A. L. & Green, D. H. *Contrib. Miner. Petrol.* **73**, 287–310 (1980).
- Kamb, W. B. *J. geophys. Res.* **64**, 1891–1909 (1959).
- Park, J. K. *Can. J. Earth Sci.* **11**, 1012–1017 (1974).
- Ernst, R. E. in *Mafic Dykes and Emplacement Mechanisms* (eds Parker, A. J., Rickwood, P. C. & Tucker, D. H.) 231–235 (Balkema, Rotterdam, 1990).

ACKNOWLEDGEMENTS. We thank K. Buchan of the Geological Survey of Canada for field support and laboratory facilities, J. Park for the samples at 1,400 km, S. Gandhi for field assistance, W. Bleeker for information on the 2,100-km sites and the Manitoba Geological Survey for logistical support. We also thank A. LeCheminant and A. Peterson for comments on the manuscript and W. Fahrig (GSC) for his pioneering work on dyke swarms.

Earliest known Australian Tertiary mammal fauna

Henk Godthelp*, Michael Archer*, Richard Cifelli†, Suzanne J. Hand* & Coral F. Gilkeson‡

* School of Biological Science, University of New South Wales, PO Box 1, Kensington, New South Wales 2033, Australia

† Oklahoma Museum of Natural History, University of Oklahoma, 1335 Asp Avenue, Norman, Oklahoma 73019, USA

‡ Westmead Dental Clinic, Westmead Hospital, Westmead, New South Wales, Australia

REMAINS of Early Eocene vertebrates from freshwater clays near Murgon, southeastern Queensland, represent Australia's oldest marsupials, bats, non-volant placentals, frogs, madtsoiid snakes, trionychid turtles¹ and birds. Radiometric dating of illites forming part of the matrix of the mammal-bearing zone has given a minimum age estimate of $54.6 \pm 0.05 \times 10^6$ years, which is roughly twice as old as any marsupials previously known from Australia² and well before the 38 million year (Myr) separation of Australia from Antarctica/South America³. All marsupials so far known from the Tingamarra Local Fauna are more derived (being dilambdodont) than peradectids. None of them is clearly a member of a previously known Australian family, but some could be uniquely plesiomorphic dasyuroids or perameloids. Another is autapomorphically specialized and indicative of at least partial isolation of the Australian portion of Gondwana. Here we report on the discovery of a tooth of the earliest non-volant placental known from Australia, *Tingamarra porterorum* gen. et sp. nov., which seems to be a condylarth-like placental mammal. The presence of non-volant placentals in the Early Tertiary of Australia challenges a common presumption that marsupials dominated Australia's therian assemblages because of failure of such placentals to reach Australia before the Late Tertiary.

A diverse assemblage of vertebrate fossils, the Tingamarra Local Fauna, is being recovered from green clays near Murgon, southeastern Queensland. Previously, only turtles and crocodiles were known from this deposit¹. Previous age estimates range from Miocene to Early Tertiary age^{1,4-6}. Potassium/argon dating of a superpositional basalt has revealed a minimum age estimate of 29.0 ± 0.2 Myr. There is another basalt beneath the deposit but it has failed to produce samples suitable for radiometric dating. The deposit outcrops in two localized areas within 1 km of each other but only one of these has produced mammals. The sediments are authigenic green illite and smectite clays with irregular dolomitized horizons and fine sandy lenses. A sample of the clay from the mammal-bearing locality and horizon dated by potassium/argon ratio revealed a minimum age of 54.6 ± 0.05 Myr. Amdel concluded: "The poor crystallinity detected by the XRD [X-ray diffractometry] suggests that illite crystallization had not progressed for too long. The age determined, therefore, is probably a minimum indication of the time of illite formation, assuming little or no loss of the radiogenic argon from the poorly crystalline illite". Procedures for K/Ar dating of authigenic illite-smectite clays have recently been reviewed⁷ and can produce reliable minimum age determinations. The Early Tertiary age of the Tingamarra deposit is corroborated by elements in the faunal assemblage including a palaeochiropterygoid-like bat. Taxonomic nomenclature of marsupials follows Aplin and Archer⁸.

Class Mammalia

Subclass Theria

Infraclass Placentalia

Order? Condylarthra

Tingamarra porterorum gen. et sp. nov.

Etymology. *Tingamarra* is an Aboriginal word meaning 'permanent water'. It is used here in reference to a local legend and the presumed palaeoecological situation. The species name is in honour of J. and M. Porter who kindly allowed us to excavate their back yard.

Holotype. QMF20564, isolated right lower molar, probably an M2 or M3.

Type locality and local fauna. Latitude 26° S; Longitude 152° E. Main Quarry, property J. & M. Porter, Boat Mountain area, Murgon, southeastern Queensland; Tingamarra Local Fauna.

Type species. *Tingamarra porterorum* sp. nov.

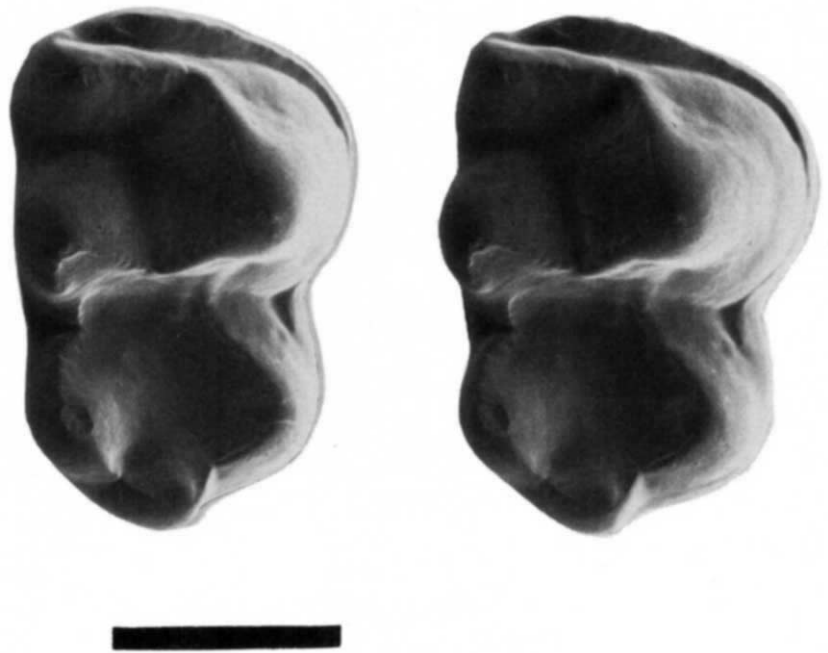
Diagnosis. The species of *Tingamarra* differ from all marsupials in having as a combination of features a non-twinned hypoconulid and entoconid, lack of a well developed buccal postcingulid and anteroposteriorly compressed trigonid. They differ from most placentals except some archaic condylarths in also having as a combination of features a broadly open trigonid and a lingually-situated paraconid that is also well anterior to the protoconid. They differ from known condylarths in having a relatively well-developed paraconid, very large entoconid and relatively less anteroposteriorly compressed trigonid.

In dental macrostructure, *Tingamarra porterorum* is unlike any marsupial known but similar to some archaic placentals. Lack of a lingually situated hypoconulid or well developed buccal postcingulid (marsupial synapomorphies⁹) are conspicuously absent in this tooth and indicate that the tooth is most probably placental.

This hypothesis is corroborated by enamel ultrastructure which is similar to that of placental mammals and unlike that of any known marsupial. The naturally worn surface of the protoconid was examined by scanning electronic microscopy (SEM) and confocal laser scanning microscopy (CLSM). Although examination by SEM was limited to this area, it was evident that prisms were present in at least the inner two-thirds of the enamel. CLSM confirmed the presence of prisms and, from examination of larger areas of enamel, permitted better determination of the prism pattern. The prisms are horseshoe-shaped, roughly 3 µm in diameter, with the open end of the horseshoe facing away from the enamel-dentine junction. A minor planar discontinuity or seam¹⁰⁻¹³ extends from within the body of the prism, through the open end of the horseshoe, into the inter-prism. The precise limits of the seams could not be determined. No tubules were observed away from the enamel-dentine junction and the prism-packing pattern was variously pattern 2 or pattern 3¹⁴. Areas of pattern 2 enamel have well separated prisms and large areas of inter-prism but no clearly defined inter-row sheet. The pattern 2 arrangement is similar to that in placentals such as microchiropterans¹⁰ and unlike that observed in marsupials^{15,16}. Overall, prism-packing pattern and the absence of tubules is more characteristic of some placental but no marsupial enamel known.

Compared with the earliest known placentals^{17,18}, the *Tingamarra* tooth is plesiomorphic in retaining a broadly open trigonid with a lingual paraconid that is also situated well anterior to the protoconid¹⁹. Similarly, in many Late Cretaceous placentals this cusp is more mesially situated²⁰. But its position and size are similar to that seen in a few archaic Late Cretaceous to Early Tertiary ungulate or ungulate-like mammals placed in the order 'Condylarthra'²¹. Examples include *Protungulatum* and *Oxyclaenus* which are generally placed in the Arctocyonidae^{22,23} and undescribed condylarths of possible Coniacian (or Late Campanian/Maastrichtian) age from Soviet Middle Asia²⁴. These share with the *Tingamarra* tooth a similar degree of reduction in trigonid height and talonid broadening yet retain a plesiomorphically open trigonid with the lingual paraconid situated well anterior to the protoconid and the paracristid relatively unreduced. By contrast, most other archaic ungulates (for example hyposodontids, mioclaenids, peripitychids, phenacodontids and didolodontids) exhibit a more reduced and

FIG. 1 Stereo occlusal scanning electron microphotograph of the Holotype (QMF20564, a lower right molar) of the Early Eocene placental *Tingamarra porterorum* from Murgon, Queensland, Australia. Scale bar, 1 mm.



posteriorly situated paraconid. The Palaeocene South American ?condylarth *Perutherium* has been argued to be a basal notoungulate²⁵. Paraconid morphology of all known South American placentals (including condylarths of Palaeocene age), in its relatively reduced state and more mesially situated position²⁶, is more derived than that of *Tingamarra porterorum*. But common ancestry of the *Tingamarra* taxon with one or more of these relatively derived placental groups is not prohibited by its relatively plesiomorphic structure. The large entoconid of the *Tingamarra* tooth is a conspicuous, apparent autapomorphy that invites wider comparison. It is also possible that the *Tingamarra* taxon is related to plesiomorphic pantodonts such as *Alcidedibignya* from the Palaeocene of Bolivia²⁵. But similarities to these archaic placentals are not as striking as those shared with plesiomorphic Cretaceous/Tertiary condylarths.

The *Tingamarra* Local Fauna is the only Australian marsupial-bearing assemblage that predates the separation of Australia from the rest of Gondwana³. Its marsupials are twice as old as any previously known from Australia². So far, a minimum of 10 taxa are represented by 50 teeth. Most taxa are essentially plesiomorphic but a few are relatively derived. The affinities of all *Tingamarra* mammal taxa are ambiguous. None can confidently be assigned to any previously known family. Because marsupials so far recovered are all dilambdodont (and thus unlike Cretaceous/Tertiary peradectids, stagodontids or other plesiomorphic rectodont marsupials⁸), there are broad resemblances to didelphimorphians, dasyuromorphians and peramelemorphians. Although few exhibit family-specific synapomorphies, one resembles a family of marsupials from the Casamayoran (Early Eocene) of South America²⁷ and another the Oligo-Miocene Australian ?dasyurid genus *Ankotarinja*²⁸.

Considered as a whole, the *Tingamarra* assemblage is unlike any of the previously oldest-known (Oligo-Miocene) faunas of Australia^{2,29} most notably in the dominance of plesiomorphic insectivorous marsupials. Compared with South American faunas, the *Tingamarra* mammals more closely resemble Early Tertiary assemblages such as the Palaeocene Tiupampian Local Fauna of Bolivia²⁵ than the Late Cretaceous Los Alamos Local Fauna of Patagonia³⁰. The as yet poorly known mammal assemblage of the Middle Eocene La Meseta Local Fauna of Antarctica^{31,32} contains marsupial and edentate groups shared with South America but none of the family-level groups currently recognized in the *Tingamarra* assemblage.

Tingamarra porterorum is the earliest non-volant placental mammal known from Australia. Its presence in the *Tingamarra* Local Fauna demonstrates that Cainozoic dominance of Australia by marsupials should not be attributed to the failure of such placentals to reach this continent in the Early Tertiary. As in the case of South America and probably Antarctica, both placentals and marsupials were present at the beginning of the Cainozoic but, for whatever serendipitous reasons, early placentals failed to survive in Australia. Non-volant placentals did not re-establish in Australia until the arrival of Eurasian murids sometime in the Late Miocene or earliest Pliocene^{33,34}.

Bonaparte³⁰ suggests that Cretaceous mammal faunas of the southern continents were mainly isolated from those of the northern continents. The highly distinctive Late Cretaceous Los Alamos mammals³⁰ and apparent confinement of Cretaceous monotremes to the Southern Hemisphere³⁵ supports this hypothesis. The presence of tiny condylarth-like placentals in the Palaeocene of South America²⁵ and Early Eocene of Australia, ornithorhynchid-like monotremes in the Palaeocene of South America³⁶ and Cainozoic of Australia and freshwater ceratodontid lungfish³⁷ (among other shared vertebrate groups) in the Palaeocene/Eocene of both continents also suggests that within the Southern Hemisphere the Australian portion of Gondwana was connected, albeit intermittently by filter barriers, to the rest of Gondwana between Late Cretaceous and Early Eocene time.

A shallow, low-energy aquatic environment is indicated by the *Tingamarra* sediments, the lack of abrasive damage to the fossils and the composition of the fauna. The nature of the surrounding vegetation and climate is unclear but there are no undoubted folivores among mammals so far recovered. □

Received 28 October 1991; accepted 7 February 1992.

1. Gaffney, E. S. & Bartholomai, A. *J. Paleont.* **53**, 657–673 (1982).
2. Archer, M., Godthelp, H., Hand, S. J. & Megirian, D. *Aust. Zool.* **25**, 29–65 (1989).
3. Veevers, J. J. (ed.) *Phanerozoic Earth History of Australia* (Clarendon, Oxford, 1984).
4. Godthelp, H., Archer, M. & Hand, S. J. *Abstracts Conf. Aust. Vert. Evol. Palaeont. System. Sydney* (1989).
5. Murphy, P. R., Schwarzbock, H., Cranfield, L. C., Withnall, I. W. & Murray, C. G. *Rep. geol. Surv. Qld* **96**, 1–157 (1976).
6. Molnar, R. *Mem. Qd Mus.* **20**, 657–673 (1982).
7. Mossman, J.-R. *Clay Minerals* **26**, 189–198 (1991).
8. Aplin, K. & Archer, M. in *Possums and Opossums: Studies in Evolution* (ed. Archer, M.) xv–lxxii (Surrey Beatty and The Royal Zoological Society of New South Wales, Sydney, 1987).
9. Cifelli, R. L. *Vert. Paleont.* **10**, 320–331 (1990).
10. Lester, K. S. & Hand, S. J. *Scanning Microsc.* **1**, 421–436 (1987).

11. Lester, K. S. & Boyde, A. *Adv. Dent. Res.* **1**, 181–190 (1987).
12. Lester, K. S. *Scanning Microsc.* **3**, 639–644 (1989).
13. Lester, K. S. & von Koenigswald, W. *Scanning Microsc.* **3**, 645–663 (1989).
14. Boyde, A. thesis, Univ. London (1964).
15. Lester, K. S., Boyde, A., Gilkeson, C. & Archer, M. *Scanning Microsc.* **1**, 401–420 (1987).
16. Gilkeson, C. F. & Lester, K. S. *Scanning Microsc.* **3**, 177–191 (1989).
17. Nessov, L. A. *Vest. Leningrad. Univ. geol. Geogr.* **17**, 8–18 (1985).
18. Kielan-Jaworowska, Z. & Dashzeveg, D. *Zool. Scripta* **18**, 347–355 (1989).
19. Clemens, W. A. & Lillegraven, J. A. *Contrib. Geol. Univ. Wyoming spec. pap.* **3**, 55–85 (1986).
20. Fox, R. C. *Carnegie Mus. nat. Hist. spec. pap.* **9**, 9–20 (1984).
21. Prothero, D. R., Manning, E. M. & Fischer, M. in *The Phylogeny and Classification of the Tetrapods Vol. 2: Mammals* (ed. Benton, M. J.) 201–234 (Clarendon, Oxford, 1988).
22. Sloan, R. E. & Van Valen, L. *Science* **148**, 220–227 (1965).
23. Cifelli, R. L. *Am. Mus. Novit.* **2772**, 1–49 (1983).
24. Nessov, L. A. & Kielan-Jaworowska, Z. in *5th Symp. Mesozoic Terrestrial Ecosystems Biota* 51–52 (1991).
25. Marshall, L. G. & de Muizon, C. *Nat. Geog. Res.* **4**, 23–55 (1988).
26. Cifelli, R. L. in *Mammal Phylogeny* (eds Szalay, F. S., Novacek, M. J. & McKenna, M. C.) (Springer, New York, in the press).
27. Cifelli, R. L. *Am. Mus. Novit.* **2820**, 1–26 (1985).
28. Archer, M. *Trans. R. Soc. S. Aust.* **100**, 53–73 (1976).
29. Woodburne, M. O. *et al. Spec. Publ. S. Aust. Dept. Mines and Energy* **5**, 347–363 (1985).
30. Bonaparte, J. F. *Nat. Geog. Res.* **6**, 63–93 (1990).
31. Woodburne, M. O. & Zinsmeister, W. J. *J. Paleont.* **58**, 913–948 (1984).
32. Pascual, R., Reguero, M. A., Scillato-Yane, G. J., Tonni, E. P. & Vizcaino, S. F. *Abstracts ICSEB* (1990).
33. Archer, M. & Bartholomai, A. *Alcheringa* **2**, 1–19 (1978).
34. Godthelp, H. *Mem. Qd Mus.* **28**, 171–173 (1989).
35. Archer, M., Flannery, T. F., Ritchie, A. & Molnar, R. E. *Nature* **318**, 363–366 (1985).
36. Pascual, R. *et al. Nature* (in the press).
37. Pascual, R. & Bondesio, P. in *Actus VI Congr. Geol. Arg.* **1**, 565–577 (Buenos Aires, Argentina, 1975).

ACKNOWLEDGEMENTS. Work at Murgon has been supported by the Australian Research Council, the Department of Arts, Sport, Environment, Tourism and Territories, National Estate Programme Grants (Queensland), the Queensland Museum and the University of New South Wales. The SEM photograph was taken by C.G. (Westmead Dental Clinic, Sydney).

Crystal assembly and phylogenetic evolution in heterococcoliths

Jeremy R. Young*, Jon M. Didymus†, Paul R. Bown‡, Ben Prins§ & Stephen Mann†

* Palaeontology Department, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

† School of Chemistry, University of Bath, Bath BA2 7AY, UK

‡ Department of Geological Sciences, University College, London WC1E 6BT, UK

§ Shell International Petroleum, 2501 AN The Hague, The Netherlands

COCCOLITHS, the calcite plates formed by unicellular phytoplanktonic algae (coccolithophores, phylum Prymnesiophyta)¹, are produced in enormous quantities and probably constitute the largest single carbonate sink in oceanic biogeochemical cycles. They are major sediment formers, are of great value to geologists as biostratigraphic indicators and have been extensively studied as models for biomineralization^{2,3}. We have applied the understanding of coccolith biomineralization to the fossil record⁴ and present evidence from electron and optical microscopy that there has been a conserved mechanism of crystal nucleation throughout the 230 million year history of coccolithophores. This fundamental feature of coccolith growth, which we term the V/R model, involves the assembly of a ring of single crystals with alternating orientations, radial (R) and vertical (V), and provides a powerful tool for tracing phylogenies, identifying homologous structures and rationalizing the higher taxonomy of the group. The living coccolithophorid, *Emiliania huxleyi*, seemed anomalous because only R crystals had been observed; transmission electron microscopy

of proto-coccolith rings, however, revealed relict V crystals which are overgrown by preferential development of R units in complete coccoliths. A speculative model based on a plicated macromolecular template is presented as a possible explanation for the conserved nucleation process.

Of the two basic types of coccoliths, heterococcoliths and holococcoliths^{1,5}, holococcoliths are formed of simple-shaped tessellating crystals whereas heterococcoliths are formed of a radial array of complex-shaped crystals. Heterococcoliths can have a variety of final morphologies⁵, but the most common are placoliths, consisting of two disc-like shields connected by a tube (for example, *E. huxleyi*, Fig. 1). Here we concentrate on placoliths, but we believe that our results will prove applicable to heterococcoliths in general. Studies on *E. huxleyi* and other living coccolithophores have shown that heterococcolith biomineralization is a two-stage process with a clear division between nucleation and crystal growth^{6–11} (Fig. 1). Nucleation produces a proto-coccolith ring of calcite (CaCO₃) crystals around the margin of a precursor organic base-plate^{8,9}. The crystals initially have simple shapes but closely regulated spacing and crystallographic orientation¹⁰. Growth develops these initial crystals into complex genus-characteristic ultrastructures. We have investigated key living and fossil coccolith genera and revised existing descriptions using both electron microscopy observations to identify the locus of the proto-coccolith ring and the form of the crystal units, and light-microscopy determination of crystallographic orientation^{12,13}. In addition, *E. huxleyi* proto-coccolith rings were studied by high-resolution transmission electron microscopy.

A detailed analysis has been made of the Mesozoic coccolith, *Watznaueria*⁴ (Figs 2 and 3), showing that the bulk of the coccolith is formed of a cycle of R units that produces a bright pseudo-extinction cross when the coccolith is observed in plan view. A second, minor cycle is located in the core of the tube

FIG. 1 Growth sequence of *E. huxleyi* showing seven stages of development. The sequence shows how, by accretionary growth, the initial simple crystals of the proto-coccolith ring develop into the morphologically complex crystal units of the complete coccolith, each composed of several superficially discrete elements. Each diagram shows two main (R) crystal units and adjacent interstitial (V) crystals. Inset, diagram of a complete *E. huxleyi* coccolith.

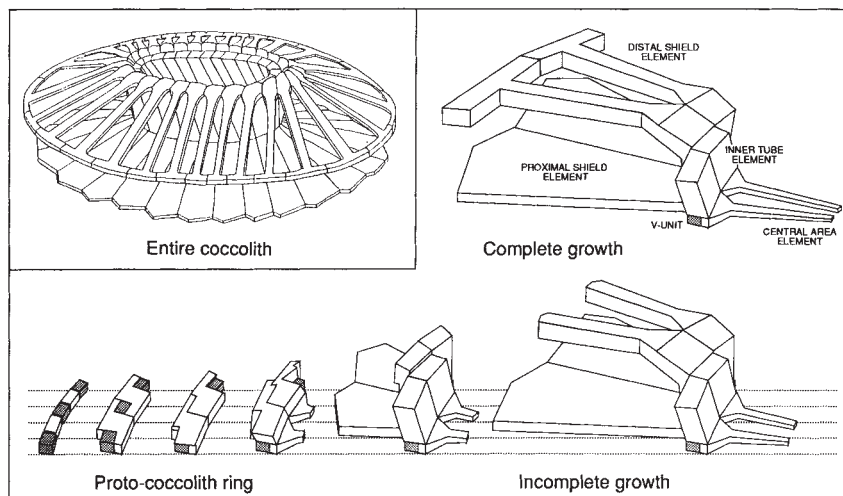
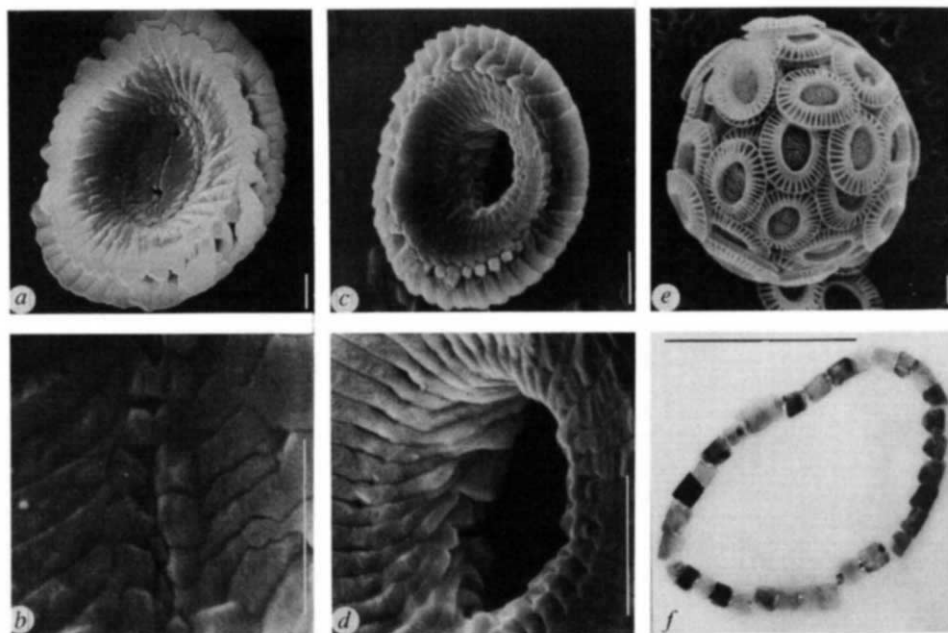


FIG. 2 Scanning electron micrographs of coccoliths, scale bars, 1 μm . *a, b*, Proximal views of *Watznaueria barnesae* (sample AG3, Cenomanian Chalk from Folkestone, England; negative numbers BM(NH) 104241 & 88068). The bulk of the coccolith is formed of R units (see Fig. 3); *a*, individual crystals can be traced from the centre of the coccolith across the proximal shield and through overgrowths, to the distal shield. The V-unit bases can be seen in a band around the central area (arrows in *b*), corresponding to the proto-coccolith ring of alternating V and R nuclei. *c, d*, proximal views of *Toweius magnicrassus* (from early Eocene (NP12) of South Atlantic, DSDP sample 549-11-3; negative numbers BM(NH) 99786 & 99785). R units form the tube cycles and lower part of the proximal shield (see Fig. 3, but note that there are no central area elements in this species). V units form the upper proximal shield and distal shield cycles, here almost completely united by overgrowth. The V unit bases are represented by a band of slits around the central area (arrows in *d*), corresponding to the proto-coccolith ring. *e*, Coccosphere and dissociated coccoliths of *E. huxleyi* (filter sample of living nannoplankton from North Atlantic surface waters, IOS 11290/2/11, courtesy of R. Jordan; negative number BM(NH) 82925). Only R units are observed. *f*, Transmission electron micrograph of *E. huxleyi* proto-coccolith ring (sample JY6/8, isolated



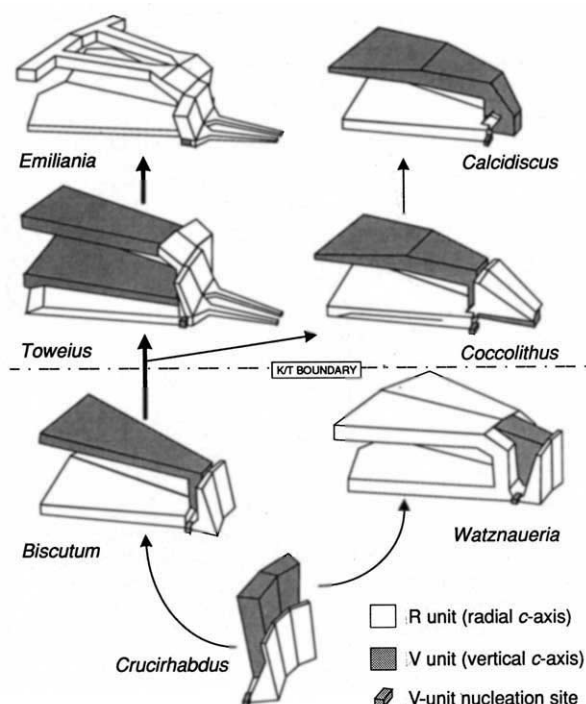
from laboratory culture of strain L, ref. 11; negative number JMD 26039). The larger crystals have radial crystallographic orientations and would develop into R units; the smaller interstitial crystals (solid arrows) are interpreted as relict V units. Some sites are absent owing to loss of the smaller crystals during sample preparation (open arrows).

and consists of peg-like V crystals that appear dark in plan view in cross-polarized light. Thus, the V and R units are oriented with the isotropic calcite *c*-axis vertical and radial to the coccolith plane, respectively. In proximal view, the two crystal units alternate around a narrow central belt (Fig. 2*b*). Scanning electron micrographs of *Watznaueria* coccoliths at different stages of growth⁴ clearly indicated that this belt was the locus of a proto-coccolith ring of crystallographically alternating nuclei.

Figure 3 summarizes our interpretations of the crystal assembly of representative placolith coccoliths and of *Crucirhab-*

dus, the most probable ancestral coccolith^{14,15}. In each case, previous observations had shown that all the elements had either vertical or sub-radial crystallographic orientations^{12,14,16}. The elements interconnect so that the apparent multi-element structures resolve into just two crystal units, R and V. Moreover, as in *Watznaueria*, the crystal units appear to originate from a belt on the proximal side around which R and V units alternate. These examples, with their immediate relatives, constitute a continuous phylogenetic series spanning the fossil record of coccoliths from the Triassic to the present, indicating that this

FIG. 3 Ultrastructure of key Triassic to Modern heterococcolith genera. Thick arrows indicate phylogenetic relationships that are well established from the fossil record; thin arrows indicate likely phylogenetic relationships with less direct evidence^{5,12,14,15}. K/T Boundary, Cretaceous/Tertiary boundary. It is suggested that the placolith form has evolved homeomorphically but that the V and R units are homologous throughout. Each diagram represents two R units, one V unit and the nucleation site of a second V unit, except *Crucirhabdus* (two V and three R units). Note that for clarity, the units are shown as more orthogonal and thinner than they actually are. Also, in most cases the illustrated genera did not evolve directly from one another but through intermediates. In addition to our observations, sources: *Emiliania*^{10,11,19,20}, with addition of V-unit nuclei; *Toweius*^{12,22}, with reinterpretation of crystallography; *Biscutum* and *Crucirhabdus*^{14,15}; *Calcidiscus*¹⁶; *Coccolithus*¹², with re-interpretation of crystallography and central area; *Watznaueria*¹⁵. Geological ranges^{5,15}. *Emiliania*, Holocene (*Reticulofenestra* and *Gephyrocapsa* have similar structures and range from Eocene to Recent); *Toweius*, Palaeocene–Eocene; *Biscutum*, Sinemurian–Maastrichtian; *Crucirhabdus*, Rhaetian–Toarcian; *Calcidiscus*, Miocene–Recent; *Coccolithus*, Palaeocene–Recent.



nucleation pattern has been stable over about 230 million years. Light-microscope observations on other heterococcolith genera indicate that most are formed of V and R units¹⁴. Detailed electron microscopy observations are necessary to test whether the crystals did indeed grow from a proto-coccolith ring of alternating nuclei, but it seems likely that this was the dominant pattern.

The ultrastructure of *E. huxleyi* and other Neogene reticulofenestrid coccoliths¹⁰ seemed anomalous because only R units had been observed. The structure of *E. huxleyi* coccoliths had been rigorously investigated^{8,11,17-21} and each complete coccolith had been thought to consist of a single cycle of crystals with radial crystallographic *c*-axes. We have now found an additional cycle of crystals in proto-coccolith rings at very early stages of growth (Fig. 2f); these crystals are rectangular in projection, 30–60-nm long and are located interstitially between the larger R units. They often seemed to be removed from the ring during sample preparation (Fig. 2f) and were overgrown by the R units at later stages of development (Fig. 1), which could explain their oversight in previous investigations. As they are extremely small, it is not possible to determine their crystallographic orientation by light microscopy or electron diffraction and it has not been possible to obtain lattice images from them, but it is clear from the transmission electron micrographs that they are spatially discrete from the R units. Our proposed V/R model provides an immediate interpretation of these small crystals as relict V unit nuclei.

The phylogenetic origins of *Emiliania* are well established from the fossil record (Fig. 3)^{5,12,22}. The genus is believed to have evolved from *Toweius* through *Reticulofenestra* and *Gephyrocapsa*, both of which have *Emiliania*-type ultrastructures¹⁰. *Toweius* (Figs 2, 3) has a well developed V unit but also a complex R unit which is similar to that of *Emiliania* except that the distal shield element is very short. Thus, evolution could have occurred by retraction of the V unit and expansion of the R unit, rather than by rotation of crystallographic axes as assumed implicitly in previous work^{12,22}. The critical evolutionary step from *Toweius* to *Reticulofenestra* occurred in the Early Eocene^{5,12,22}, so it seems that non-functional V units have continued to be formed through a further 50 million years of

evolution. The conservation of this crystal nucleation pattern over long periods of time suggests that it is regulated by a fundamental and relatively simple mechanism. Electron diffraction studies have identified the calcite {110} face as the probable nucleation surface in *E. huxleyi*¹⁹, and we have confirmed this with similar studies on immature crystals within proto-coccolith rings (data not shown). It has been suggested that preferential nucleation of the R-unit {110} face is due to specific structural and stereochemical interactions between the calcite nuclei and an underlying macromolecular surface which displays an array of binding sites that mirror the calcite structure¹⁹.

V- and R-crystal orientations might be produced by alternating two macromolecular templates; a more economical hypothesis suggests they are produced by folding of a single macromolecular template. Simple models show that the change in orientation between adjacent nuclei can only be achieved with a single template by using a double fold. Repetition of this fold can produce a belt of regularly spaced alternating horizontal and vertical sections from a single linear template (Fig. 4). Moreover, the folds are intrinsically flexible and so can explain the curvature of the nucleation belt around the proto-coccolith ring. We therefore speculate that a plicated macromolecular sheet could be responsible for the periodic and site-specific nucleation of oriented calcite crystals observed in heterococcoliths. A macromolecular template mirroring a particular calcite face could not easily be evolved to mirror another; this evolutionary inertia and the proposed folding make the model structure inherently stable, in agreement with the observed stability of the nucleation pattern during coccolith evolution.

A possible candidate for the molecular template is a complex acidic polysaccharide closely associated with coccolith formation²³ but its biochemical investigation is beyond the scope of this work. Clearly, matrix biosynthesis is under strict genetic control and resilient to modification. By comparison, the morphology of R and V units has been subject to evolutionary change throughout heterococcolith history, even among placoliths. The proposed V/R model of conserved nucleation implies that, throughout this change, crystallographic orientation remains a reliable indicator of homology, and is therefore crucial to understanding coccolith phylogeny. □

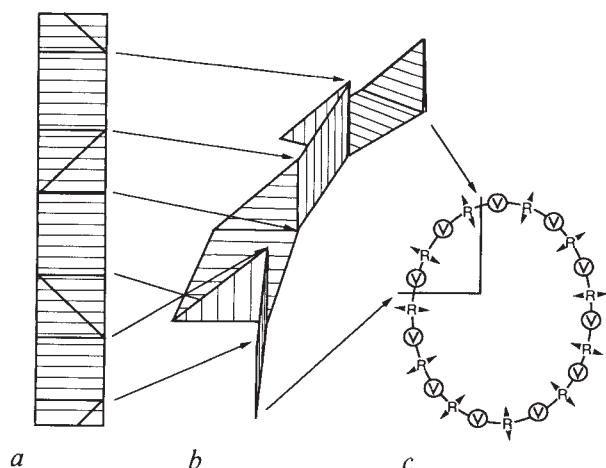


FIG. 4 Speculative model for R- and V-unit nucleation by a plicated macromolecular template. a, Planar ribbon-like macromolecular substrate with periodic binding sites for oriented calcite nucleation; thin lines represent the orientation of the crystallographic *c*-axis of the nuclei. b, Plicated belt of the linear template showing how folding can produce alternating segments with vertical and horizontal *c*-axis orientations. c, Schematic representation of alternating R and V units around a proto-coccolith ring resulting from continuation of a plicated belt around the organic base-plate precursor to the mineralized coccolith.

Received 29 October 1991; accepted 21 January 1992.

- Green, J. G., Perch-Nielsen, K. & Westbroek, P. in *Handbook of Protozoa* (eds Margulis, L., Corliss, J. O., Melkonian, M. & Chapman, D. J.) 293–317 (Jones and Bartlett, Woods Hole, 1989).
- Lowenstam, H. A. & Weiner, S. *On Biomineralization* (Oxford University Press, 1989).
- Simkiss, K. & Wilbur, K. M. *Biomineralization: Cell Biology and Mineral Deposition* (Academic Press, San Diego, 1989).
- Young, J. R. & Bown, P. R. *Palaeontology* **34**, 843–850 (1991).
- Perch-Nielsen, K. in *Plankton Stratigraphy* (eds Bolli, H. M., Saunders, J. D. & Perch-Nielsen, K.) 329–554 (Cambridge University Press, 1985).
- Wilbur, K. M. & Watabe, N. *Ann. N.Y. Acad. Sci.* **103**, 82–112 (1963).
- Klavens, D. *Protistologica* **12**, 217–224 (1976).
- Westbroek, P. et al. *Phil. Trans. R. Soc.* **B304**, 435–444 (1984).
- Manton, I. & Leedale, G. F. *J. mar. Biol. Ass. U.K.* **49**, 1–16 (1969).
- Young, J. R. in *Nannofossils and their Applications* (eds Crux, J. & Heck, S. E. van) 1–20 (Ellis Horwood, Chichester, 1989).
- Young, J. R. & Westbroek, P. *Mar. Micropalaeont.* **18**, 5–23 (1991).
- Romein, A. J. T. *Utrecht Micropalaeont. Bull.* **22**, 1–231 (1979).
- Moskowitz, S. & Osmond, K. in *Nannofossils and their Applications* (eds Crux, J. & Heck, S. E. van) 76–97 (Ellis Horwood, Chichester, 1989).
- Prins, B. in *Proc. 1st Conf. Planktonic Microfossils Vol. 2* (eds Brönnimann, P. & Renz, H. H.) 547–558 (Brill, Leiden, 1969).
- Bown, P. R. *Spec. Pap. Palaeont.* **38**, 1–118 (1987).
- Gartner, S. *Univ. Kansas Paleont. Conts. paper* **28**, 1–4.
- Braarud, T., Gaarder, K. R., Markali, J. & Nordli, E. *Nytt Mag. Bott.* **1**, 129–134 (1952).
- Watabe, N. *Calcif. Tissue Res.* **1**, 114–127 (1967).
- Mann, S. & Sparks, N. H. C. *Proc. R. Soc. B234*, 441–453 (1988).
- Emburg, P. R. in *Coccolith formation in Emiliania huxleyi*, 1–145 (University of Leiden Geobiochemistry Unit, 1989).
- Westbroek, P., Young, J. R. & Linschooten, K. *J. Protozool.* **36**, 368–373 (1989).
- Gallagher, L. in *Nannofossils and their Applications* (eds Crux, J. & Heck, S. E. van) 41–75 (Ellis Horwood, Chichester, 1989).
- Vrind-de Jong, E. W. de et al. in *Biomineralization in Lower Plants and Animals* (eds Leadbeater, B. S. C. & Riding, R.) 204–220 (Oxford University Press, 1986).

ACKNOWLEDGEMENTS. We thank P. Westbroek and his group for collaboration, L. Gallagher, M. Girgis, S. Rabinathan and O. Varol for insights on the structure of particular coccoliths P. Pearson, J. Whittaker and S. Wyers for comments. This work was supported by the NERC, the SERC, Shell International Petroleum and Clyde Petroleum.

Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*

David J. Begun & Charles F. Aquadro

Section of Genetics and Development, Biotechnology Building,
Cornell University, Ithaca, New York 14853-2703, USA

TWO genomic regions with unusually low recombination rates in *Drosophila melanogaster* have normal levels of divergence but greatly reduced nucleotide diversity^{1,2}, apparently resulting from the fixation of advantageous mutations and the associated hitch-hiking effect^{3,4}. Here we show that for 20 gene regions from across the genome, the amount of nucleotide diversity in natural populations of *D. melanogaster* is positively correlated with the regional rate of recombination. This cannot be explained by variation in mutation rates and/or functional constraint, because we observe no correlation between recombination rates and DNA sequence divergence between *D. melanogaster* and its sibling species, *D. simulans*. We suggest that the correlation may result from genetic hitch-hiking associated with the fixation of advantageous mutants. Hitch-hiking thus seems to occur over a large fraction of the *Drosophila* genome and may constitute a major constraint on levels of genetic variation in nature.

Table 1 summarizes levels of DNA variation and interspecific divergence (between *D. melanogaster* and *D. simulans*) where available. These estimates of DNA polymorphism are derived from restriction site surveys with one exception (*cubitus interruptus*) and therefore are estimates of average levels of variation over 13 to 65 kilobases (kb) from each gene region. To explore the relationship between levels of DNA sequence variation and recombination rates we compared estimates of nucleotide diversity (π)⁵ and the coefficient of exchange⁶, a measure of recombination rate per physical distance.

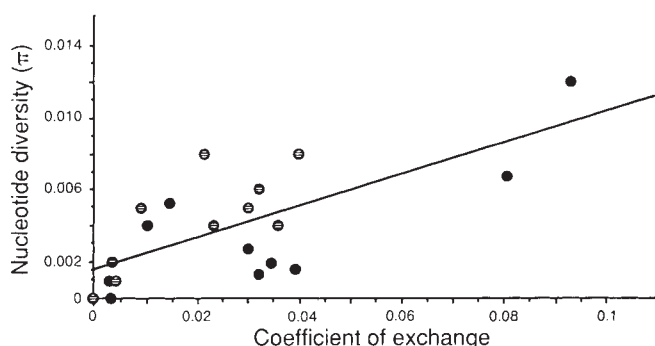


FIG. 1 Scatterplot of nucleotide diversity (π) versus coefficient of exchange in *D. melanogaster*. Autosomal and X-linked genes are represented by hatched and closed circles, respectively. To make autosomal and X-linked genes directly comparable we made the simplifying assumption of equal numbers of males and females. Then, under neutrality, nucleotide heterozygosity is an estimate of $3N\mu$ for X-linked genes and $4N\mu$ for autosomal genes (N is the effective population size and μ is the neutral mutation rate). Therefore, before doing regression or correlation analyses, we multiplied estimates of π from X-linked gene regions by four-thirds. The recombination rates estimated by the coefficient of exchange are for females. An X-linked gene region spends two-thirds of the time in females (where it can recombine) and only one-third of the time in males (where it cannot recombine), whereas an autosome spends half its time in females and half in males. Therefore, we multiplied the coefficient of exchange for autosomal genes and X-linked regions by one-half and two-thirds, respectively. Regression line is indicated by a solid line.

TABLE 1 Coefficients of exchange and nucleotide heterozygosities in *D. melanogaster* and divergence with *D. simulans*

Gene region	Coefficient of exchange	π	Divergence	Reference
Chromosome I (X)				
<i>yellow-achaete</i> (<i>y, ac</i>) phosphoglucuronate dehydrogenase gene (<i>Pgd</i>)	0.0045	0.001	0.054	1
<i>zeste-eko</i> (<i>z, tko</i>)	0.0154	0.003	0.029	1
<i>period</i> (<i>per</i>)	0.0222	0.004	—	12
<i>white</i> (<i>w</i>)	0.0520	0.001	0.050	1
<i>Notch</i> (<i>N</i>)	0.1400	0.009	—	13
<i>vermillion</i> (<i>v</i>)	0.1212	0.005	—	14
	0.0590	0.001	0.047	(D.J.B. and C.F.A., unpublished results)
<i>forked</i> (<i>f</i>)	0.0455	0.002	—	15
glucose-6-phosphate dehydrogenase gene (<i>Zw</i>)	0.0485	0.001	—	16
<i>suppressor of forked</i> (<i>su(f)</i>)	0.0050	0.000	—	15
Chromosome II				
sn-glycerol 3-phosphate dehydrogenase gene (<i>Gpdh</i>)	0.0800	0.008	—	17
alcohol dehydrogenase gene (<i>Adh</i>)	0.0647	0.006	0.045	18, 19
DOPA decarboxylase gene (<i>Ddc</i>)	0.0184	0.005	—	(C.F.A. et al., unpublished results)
amylase gene (<i>Amy</i>)	0.0435	0.008	—	20
<i>Punch</i> (<i>Pu</i>)	0.0718	0.004	—	17
Chromosome III				
esterase-6 gene (<i>Est-6</i>)	0.0604	0.005	—	21
metallothionein-A gene (<i>MtnA</i>)	0.0083	0.001	0.072	22
heat-shock protein-70A gene (<i>Hsp70A</i>)	0.0069	0.002	0.023	23, 24
<i>rosy</i> (<i>ry</i>)	0.0471	0.003	0.050	25
Chromosome IV				
<i>cubitus interruptus</i> Dominant (<i>c^D</i>)	0*	0.000	0.050	2

Nucleotide diversity (π) is the average pairwise difference for all pairs of sequences drawn at random from a population, and can be thought of as heterozygosity per nucleotide⁵. The coefficient of exchange for a gene region was calculated by selecting two genetically defined loci²⁶ that flank the region of interest and dividing the distance in map units between the flanking loci by the number of polytene bands between the loci⁶. The number of polytene bands between loci was determined from Bridge's maps²⁷. An important assumption underlying the use of this metric as an index of recombination rate is that over large stretches of the genome (for example, 20 to 40 polytene bands), the average amount of DNA per polytene band is roughly similar between regions. Available data suggest that this is a reasonable assumption for at least much of the *Drosophila* genome^{6,28,29}.

* The recombination rate on the fourth chromosome is effectively zero³⁰.

Figure 1 is a scatterplot of π versus the coefficient of exchange for 20 genes in *D. melanogaster*. It is apparent that levels of nucleotide diversity increase as rates of recombination increase. Variation in recombination rates explains a large fraction of the variation in nucleotide diversity and the null hypothesis that the slope is zero is rejected with high probability ($F_s = 16.8$, $P = 0.0007$). The non-parametric Spearman and Kendall regression tests are also significantly different from zero (Spearman's $D = 544$, $P < 0.01$; Kendall's $\tau = 0.437$, $P < 0.01$). The same conclusion is reached when data from the *white* region (which has a particularly high level of variation in *D. melanogaster*) is excluded from the analysis ($F_s = 5.8$, $P = 0.03$; Spearman's $D = 544$, $0.02 < P < 0.05$; Kendall's $\tau = 0.374$, $0.02 < P < 0.05$).

One hypothesis to explain this trend is that gene regions in areas of reduced recombination have lower neutral mutation rates. Perhaps recombination itself is mutagenic. If this were true, then under a neutral model these gene regions should also be less diverged between species than gene regions in areas of

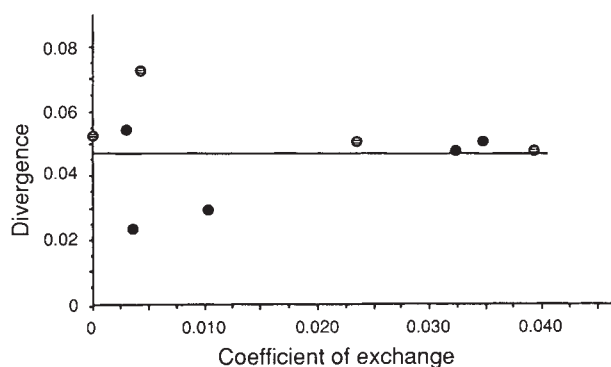


FIG. 2. Scatterplot of sequence divergence between *D. melanogaster* and *D. simulans* versus coefficient of exchange in *D. melanogaster*. Autosomal and X-linked genes are represented by hatched and closed circles, respectively. Coefficients of exchange for X-linked and autosomal regions are modified as described in Fig. 1 legend. Regression line is indicated by a solid line.

greater recombination rates⁷. In Fig. 2 we show a plot of divergence between *D. melanogaster* and *D. simulans* versus the coefficient of exchange, including all gene regions for which we have estimates of divergence. Clearly, estimates of divergence from a larger number of gene regions are desirable. Nevertheless, the lack of a significant positive regression coefficient ($F_1 = 0.001$, $P = 0.983$) with the available data argues against the hypothesis that gene regions in areas of low recombination rates have, on average, lower substitution rates.

Theoretical results show that at the time of fixation of a neutral variant, the amount of linked neutral variation is reduced, and that the magnitude of the reduction depends on the recombination rate⁸. But at a random time (which is any time a genomic region in a population is sampled), the average amount of neutral nucleotide polymorphism is unaffected by the recombination rate^{8,9}. Therefore, we are unable to arrive at a satisfactory neutral explanation for the patterns seen in Figs 1 and 2.

We propose that the positive correlation between DNA variation and recombination rate results from the selective fixation of advantageous mutants over a significant portion of the genome. This correlation suggests that levels of neutral variation in many of the gene regions for which variation has been measured have been reduced by one or more hitch-hiking events. Provided that a new selectively favoured mutation goes to fixation before another advantageous mutation arises close to it, each fixation will be surrounded by a 'window' of reduced polymorphism, the relative size of which is proportional to the rate of recombination for that region of the genome⁴. Thus, where recombination rates are very low, each fixation will cause a wide window of reduced polymorphism, whereas in regions of higher recombination, the window will be proportionately smaller. Moving along a chromosome towards regions of progressively lower recombination, the windows become closer, and may begin to overlap substantially. Thus, regions of low

recombination are 'hit' by selective sweeps more often, keeping polymorphism at a lower average level. Mutations driven to fixation by meiotic drive or biased gene conversion would have similar evolutionary consequences. Hitch-hiking does not affect interspecific divergence¹⁰, consistent with the observed lack of a correlation between DNA sequence divergence and recombination rate.

McDonald and Kreitman¹¹ proposed that patterns of synonymous and non-synonymous variation at the *Adh* locus in and between three *Drosophila* species are incompatible with neutrality. They suggested that selective fixation of amino-acid polymorphisms at this locus is the best explanation for their data. Furthermore, they speculate that selective fixations occur at a large number of loci. Our results are consistent with this view. However, the number of selectively favoured nucleotides relative to the size of the genome could still be quite small⁴.

Clearly, hitch-hiking and recombination rates do not explain all of the heterogeneity in levels of variation across the *D. melanogaster* genome. Variation in mutation rate and functional constraint, as well as several different forms of selection, have roles in shaping local levels of DNA sequence variation. But the analysis presented here presents the first evidence that hitch-hiking driven by selective fixation of new mutations may constrain levels of nucleotide polymorphism over large portions of the *D. melanogaster* genome. Inference of effective population size from levels of DNA variation may be compromised by this phenomenon.

Much effort is being expended to assemble physical and genetic maps in several species. An unexpected benefit of these genome mapping projects is that it will be possible to examine whether correlations between recombination rates and levels of DNA variation are a general phenomenon in natural populations of other taxa, including humans. □

Received 19 December 1991; accepted 7 February 1992.

- Begun, D. J. & Aquadro, C. F. *Genetics* **129**, 1147-1158 (1991).
- Berry, A. J., Ajioka, J. W. & Kreitman, M. *Genetics* **129**, 1111-1117 (1991).
- Maynard Smith, J. & Haigh, J. *Genet. Res.* **23**, 23-35 (1974).
- Kaplan, N. L., Hudson, R. R. & Langley, C. H. *Genetics* **123**, 887-899 (1989).
- Nei, M. *Molecular Evolutionary Genetics* (Columbia Univ. Press, 1987).
- Lindsley, D. L. & Sandler, L. *Phil. Trans. R. Soc. B* **277**, 295-312 (1977).
- Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
- Tajima, F. *Genetics* **125**, 447-454 (1990).
- Hudson, R. R. *Theor. Populat. Biol.* **23**, 183-201 (1983).
- Birky, C. W. Jr & Walsh, J. B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6414-6418 (1988).
- McDonald, J. H. & Kreitman, M. *Nature* **351**, 652-654 (1991).
- Aguadé, M., Miyashita, N. & Langley, C. H. *Molec. Biol. Evol.* **6**, 123-130 (1988).
- Miyashita, N. & Langley, C. H. *Genetics* **120**, 199-212 (1988).
- Schaeffer, S. W., Aquadro, C. F. & Langley, C. H. *Molec. Biol. Evol.* **5**, 30-40 (1988).
- Langley, C. H. in *Population Biology of Genes and Molecules* (eds Takahata, N. & Crow, J. F.) 75-91 (Baifukan, Japan).
- Eanes, W. F., Ajioka, J. W., Hey, J. & Wesley, C. *Molec. Biol. Evol.* **6**, 384-397 (1989).
- Takano, T. S., Kusakabe, S. & Mukai, T. *Genetics* **129**, 753-761 (1991).

- Langley, C. H., Montgomery, E. & Quattlebaum, W. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5631-5635 (1982).
- Aquadro, C. F., Deese, S. F., Bland, M. M., Langley, C. H. & Laurie-Ahlberg, C. C. *Genetics* **114**, 1165-1190 (1986).
- Langley, C. H. et al. *Genetics* **119**, 619-629 (1988).
- Game, A. Y. & Oakeshott, J. G. *Genetics* **126**, 1021-1031 (1990).
- Lange, B. W., Langley, C. H. & Stephan, W. *Genetics* **126**, 921-932 (1990).
- Leigh Brown, A. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5350-5354 (1983).
- Leigh Brown, A. J. & Ish-Horowitz, D. *Nature* **290**, 677-682 (1981).
- Aquadro, C. F., Lado, K. M. & Noon, W. A. *Genetics* **119**, 875-888 (1988).
- Ashburner, M. *Drosophila Genetic Maps* (*Drosophila* Information Service **69**, 1991).
- Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster* (Carnegie Institute, Washington DC, 1967).
- Sousa, V. *Chromosome Maps of Drosophila* (CRC, Boca Raton, Florida, 1988).
- Merriam, J., Ashburner, M., Hartl, D. L. & Kafatos, F. C. *Science* **254**, 221-225 (1991).
- Hochman, B. in *Genetics and Biology of Drosophila* Vol. 1b (eds Ashburner, M. & Novitski, E.) 903-928 (Academic, New York, 1976).

ACKNOWLEDGEMENTS. We thank M. Nachman for comments and all members of our laboratory for discussion. This work was supported by the NIH and NSF.

Protein kinase C reduces Mg^{2+} block of NMDA-receptor channels as a mechanism of modulation

Li Chen & Li-Yen Mae Huang*

Marine Biomedical Institute and Department of Physiology and Biophysics, University of Texas Medical Branch, Galveston Texas 77550, USA

* To whom correspondence should be addressed

THE roles of *N*-methyl-D-aspartate (NMDA) receptors and protein kinase C (PKC) are critical in generating and maintaining a variety of sustained neuronal responses. In the nociceptive (pain-sensing) system, tissue injury or repetitive stimulation of small-diameter afferent fibres triggers a dramatic increase in discharge (wind-up) or prolonged depolarization of spinal cord neurons. This central sensitization can neither be induced nor maintained when NMDA receptor channels are blocked^{1,2}. In the trigeminal subnucleus caudalis (a centre for processing nociceptive information from the orofacial areas³), a μ -opioid receptor agonist causes a sustained increase in NMDA-activated currents by activating intracellular PKC⁴. There is also evidence that PKC enhances NMDA-receptor-mediated glutamate responses⁴⁻⁷ and regulates long-term potentiation of synaptic transmission⁸⁻¹⁴. Despite the importance of NMDA-receptors and PKC, the mechanism by which PKC alters the NMDA response has remained unclear. Here we examine the actions of intracellularly applied PKC on NMDA-activated currents in isolated trigeminal neurons. We find that PKC potentiates the NMDA response by increasing the probability of channel openings and by reducing the voltage-dependent Mg^{2+} block of NMDA-receptor channels.

Figure 1a shows the currents evoked by NMDA at various holding potentials under whole-cell voltage clamp. The peak

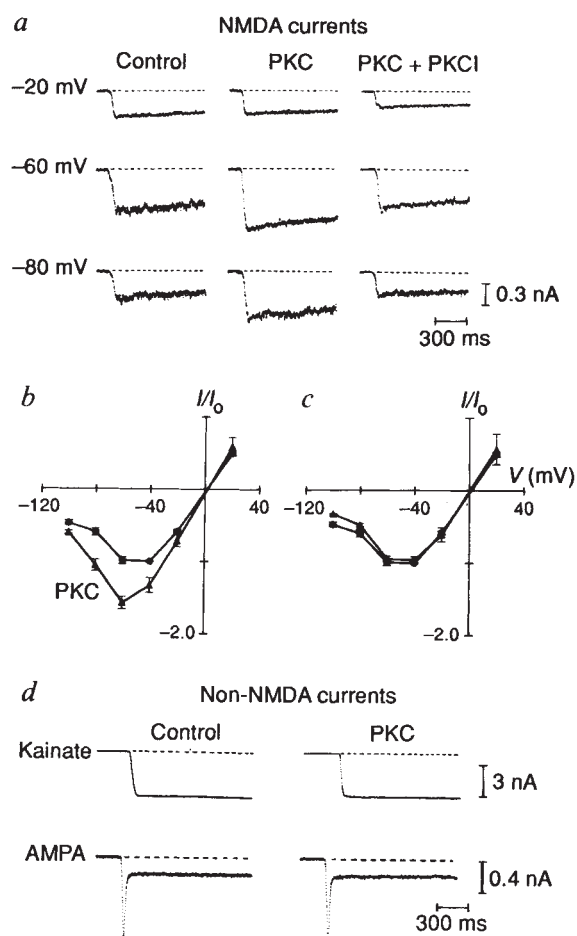
current-voltage relationship of the responses had a characteristic negative-resistance region¹⁵⁻¹⁷ (Fig. 1b). Low Mg^{2+} concentrations (30 μ M) in the external media were used in most of our experiments because NMDA-activated currents in higher Mg^{2+} (0.5–1.0 mM) at negative potentials were too small to be measured accurately. After PKC was injected into a cell, the NMDA response was gradually enhanced, reached a new steady level in 5–7 min, and stayed elevated for more than 60 min⁴. The degree of enhancement varied with membrane potential: the increase was ~80% at –80 mV but less than 20% at +20 mV (Fig. 1a, b). This effect was not seen in cells injected with heat-inactivated PKC (data not shown).

To examine the specificity of PKC action, we studied the effect of a PKC inhibitor, PKCI(19–31), a synthetic peptide that competes for the substrate recognition site of PKC^{18,19}. When PKCI alone was included in the pipette, NMDA-activated currents were decreased by 5–14%, suggesting that NMDA receptor channels might be phosphorylated by endogenous PKC. When PKC and PKCI were simultaneously perfused into the cell, the enhancing effects of PKC were blocked (Fig. 1a, c), confirming that enhancement of NMDA responses was indeed due to activation of PKC. In contrast, PKC had a minimal effect on the kainate-activated and α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-activated currents (Fig. 1d).

Additionally, we found that PKC enhanced NMDA-activated currents by 10–20% at all membrane potentials in Mg^{2+} -free solution (Fig. 2a). The apparent dissociation constant for NMDA (K_{NMDA}) did not change significantly with PKC treatment; the maximal level of NMDA-activated currents increased by 14.2% ($P < 0.05$; Fig. 2b). To confirm the direct effect of PKC on NMDA receptor channels, we determined the single-channel conductance (γ) and the probability of a channel being open (P_o) from outside-out patches (Fig. 2c). In controls, the

FIG. 1. PKC potentiates NMDA-activated currents. a, NMDA-activated currents at the indicated holding potentials were recorded in control buffer, 10 min after PKC and 5 min after PKC + PKCI are intracellularly perfused. The external buffer contained 30 μ M Mg^{2+} . b, The peak NMDA-activated currents (I) in control (●) and in PKC (▲) were plotted as a function of holding potentials (V). The data are averages from eight cells. For each cell, the currents at different potentials were normalized with respect to the current measured at –40 mV in control internal solution (I_o). Standard errors are indicated by bars; data points were connected by lines for clarity. PKC potentiates the NMDA-activated currents by 28.3% at –100 mV, 78.1% at –80 mV, 59.1% at –60 mV, 33% at –40 mV, 22.6% at –20 mV and by 18.0% at +20 mV. c, The peak current-voltage relations in control (●) and in PKC plus PKCI (▲; upper curve). Points are the average values from six cells. After simultaneous application of PKC and PKCI, NMDA-activated currents decrease slightly at negative potentials. d, At –60 mV, 200 μ M kainate (EC_{50} = 150 μ M; ref. 29) caused an inward current that did not desensitize. Addition of PKC had no effect on kainate-activated currents. AMPA at 200 μ M (EC_{50} = 80 μ M) induced a rapidly desensitizing current which was also not affected by intracellular addition of PKC (data are from two different cells).

METHODS. Isolated trigeminal neurons were prepared as described²⁹. Whole-cell currents were digitally sampled at 500 μ s per point; single channel currents obtained from outside-out patches³⁰ were sampled at 100 μ s. Both signals were filtered at 1 kHz. The external buffer contained (mM) NaCl (140), KCl (4), glucose (10), $CaCl_2$ (2.5), glycine (0.001) and the indicated amount of $MgCl_2$, HEPES (10), pH 7.35. The internal buffer contained (mM) caesium methanesulphonate (125), CsCl (15), BAPTA (11), $CaCl_2$ (3), and $MgCl_2$ (2), HEPES (20), pH 7.2. The free intracellular Ca^{2+} concentration was calculated at 100 nM (ref. 31). ATP (5 mM), leupeptin (200 μ M) and GTP (100 μ M) were added internally to prevent rundown of NMDA responses. In a few experiments, we used KF instead of caesium methanesulphonate as the chief internal electrolyte to inhibit endogenous phosphatase activity and obtained similar results. Unless specifically stated, activity rather than concentration of Mg^{2+} was used throughout our analyses. The activity coefficient for Mg^{2+} was calculated to be 0.64 (ref. 32).



main current level of single NMDA-activated channels (i) was $2.49 \text{ pA} \pm 0.04$ ($n=4$) at -70 mV . The estimated γ was 35.7 pS , similar to observed values in other central neurons²⁰⁻²³ and P_0 was $9.4 \pm 0.6\%$. After PKC addition, $i = 2.54 \text{ pA} \pm 0.03$ ($n=3$) and $P_0 = 13.7 \pm 0.6\%$. Thus, PKC had little effect on γ , but increased P_0 significantly.

We then studied the effect of PKC on the voltage-dependence of Mg^{2+} block of NMDA receptor channels in media containing different concentrations of Mg^{2+} . At each potential, NMDA-activated currents were recorded as the external Mg^{2+} concentration varied from 1 to $5,000 \text{ } \mu\text{M}$. The same experiments were repeated with PKC or PKC plus PKCI perfused into the cell. The apparent dissociation constant for Mg^{2+} block (K_{Mg}) in PKC plus PKCI internal solution was identical to that in controls (data not shown), but the K_{Mg} in PKC alone (such as K_{Mg} =

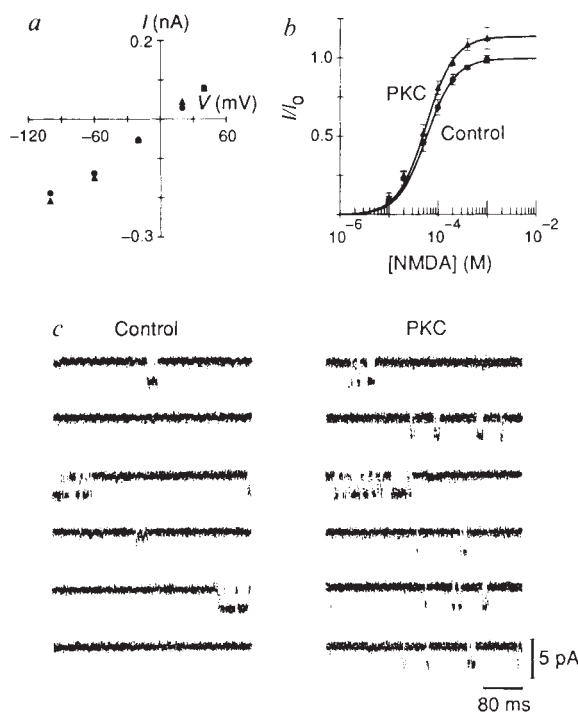


FIG. 2 PKC slightly increases NMDA responses in Mg^{2+} -free solution. *a*, The current-voltage (I - V) relationship in control buffer and with PKC. No Mg^{2+} was added externally; the basal Mg^{2+} concentration in our external solution measured with a flame photometer was $<1 \text{ } \mu\text{M}$. The changes in NMDA-activated currents before ($\bullet$) and after ($\blacktriangle$) PKC treatment were less than 20%. *b*, Dose-response curves for NMDA at -60 mV , using control buffer and internal PKC successively in the same cell. Solid lines were drawn according to the equation $I = I_0 [\text{NMDA}]^n / (K_{\text{NMDA}}^n + [\text{NMDA}]^n)$ where I_0 is the maximal current in control internal buffer, K_{NMDA} is the apparent dissociation constant for NMDA, $[\text{NMDA}]$ is the concentration of NMDA, and n is the Hill coefficient. For control, average $K_{\text{NMDA}} = 58.9 \pm 7.8 \text{ } \mu\text{M}$, ($n=6$), $I/I_0 = 1.0$ and $n=1.5$. For PKC, $K_{\text{NMDA}} = 57.6 \pm 9.2 \text{ } \mu\text{M}$ ($n=6$), $I/I_0 = 1.14$ and $n=1.5$. The change in K_{NMDA} was not significant ($P > 0.5$, Student's t -test). *c*, Single channel activities recorded from outside-out patches at -70 mV . Inward currents appear as downward deflections [NMDA $10 \text{ } \mu\text{M}$]. In control buffer, the main current level was 2.5 pA and P_0 was 0.09 . With PKC, the current level was 2.5 pA and P_0 was 0.13 . The data are from two different patches. METHODS. All chemicals were ultrapure grade to decrease trace Mg^{2+} in our solutions. The concentration of chemicals used was as follows: NMDA, $100 \text{ } \mu\text{M}$; APV, $200 \text{ } \mu\text{M}$; PKC, 0.247 U ml^{-1} (1 U is 1 nmol phosphorylation of histone per min); PKCI ($19-31$), $10 \text{ } \mu\text{M}$; diolefin $2 \text{ } \mu\text{g ml}^{-1}$. PKC and PKCI were freshly prepared before each experiment and added directly to the internal solution. To boost PKC activity, intracellular Ca^{2+} was buffered at a high level (100 nM) and diolefin was included in the internal perfusate. Under these conditions, PKC activity, estimated from a mixed micelle assay³³, was $1.7 \text{ } \mu\text{mol per min per mg protein}$. Amino acids were delivered to the recorded cells with a fast perfusion method²⁹. PKC and PKCI were internally applied through a plastic tube inserted in the patch pipet⁴. The least χ^2 method was used to fit theoretical equations to the experimental data.

$110.0 \text{ } \mu\text{M}$ at -60 mV) was 3–4 times larger than K_{Mg} in control buffer ($K_{\text{Mg}} = 27.3 \text{ } \mu\text{M}$; Fig. 3*a, b*). We assumed that the binding site for Mg^{2+} is located inside the channel, which is blocked when Mg^{2+} binds. K_{Mg} is then an exponential function of voltage such that $K_{\text{Mg}} = K_{\text{Mg}0} \exp(zF\delta V/RT)$; the binding rate of Mg^{2+} depends on the location of the binding site, the electrical distance, δ , in the membrane field, (see Fig. 3 legend for details). The observed voltage dependence of K_{Mg} was well described by this function (lines in Fig. 3*b*). In PKC-treated cells, δ increased by $\sim 11\%$; $K_{\text{Mg}0}$, the dissociation constant at 0 mV , increased 4.8 times. In contrast to the large blocking effect of external Mg^{2+} on inward NMDA-activated currents, intracellular Mg^{2+} had relatively little effect on the outward NMDA currents in our cells. At $+80 \text{ mV}$, 2.5 mM free internal Mg^{2+} blocked the outward currents by about 14%. PKC did not alter the intracellular Mg^{2+} block of NMDA-activated currents (data not shown).

To understand better the physiological significance of the increases in $K_{\text{Mg}0}$ and δ , we plotted the relative conductance (g/g_0) in $30 \text{ } \mu\text{M}$ Mg^{2+} as a function of voltage (Fig. 4*a*). Theoretical conductance voltage curves, using the estimated values of $K_{\text{Mg}0}$ and δ are also given. In controls, the midpoint of the curve is located at -62.7 mV . After PKC application, the midpoint of the curve shifted by -18.5 mV to -81.2 mV . Under physiological conditions, the free Mg^{2+} level in the extracellular space is around 0.5 mM ²⁴. The midpoint of the relative conductance voltage curve in 0.5 mM Mg^{2+} after addition of PKC, shifted -23.3 mV from -14.3 mV to -37.6 mV (Fig. 4*b*). Thus, PKC causes NMDA receptor channels to activate at more hyperpolarized potentials.

This novel way of modulating NMDA-receptor channels by PKC has two consequences. First, synaptic release of an excitatory amino acid such as glutamate would evoke a larger and longer excitatory postsynaptic potential in cells with elevated PKC, because more NMDA-receptor channels are activated at negative potentials and the NMDA currents have slow kinetics^{25,26}. Another consequence is that with more NMDA channels opened by PKC, more Ca^{2+} ions would flow into the cell^{27,28}, raising the intracellular Ca^{2+} concentration and so further increasing PKC activity. This positive feedback loop ensures sufficient protein kinase C is activated, bringing about

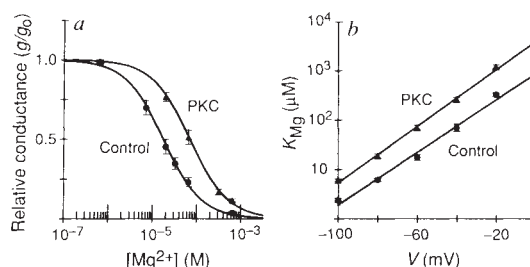


FIG. 3 PKC decreases the affinity of Mg^{2+} for NMDA-receptor channels. *a*, NMDA-activated currents recorded in different Mg^{2+} solutions. Dose response curve before ($\bullet$) and after ($\blacktriangle$) perfusion of PKC into the cells. Data points are average conductance obtained from 8 cells. Solid lines were drawn according to the equation $g/g_0 = 1 - [\text{Mg}^{2+}] / (K_{\text{Mg}} + [\text{Mg}^{2+}])$, where g_0 is the maximal conductance in Mg^{2+} free solution, K_{Mg} is the apparent dissociation constant for Mg^{2+} (activity value is used in the figure) and $[\text{Mg}^{2+}]$ is the activity of Mg^{2+} . At -60 mV , K_{Mg} was $27.3 \text{ } \mu\text{M}$ ($a_{\text{K}_{\text{Mg}}} = 17.5 \text{ } \mu\text{M}$) in control cells and $110.0 \text{ } \mu\text{M}$ ($a_{\text{K}_{\text{Mg}}} = 70.2 \text{ } \mu\text{M}$) in PKC-treated cells. *b*, The voltage-dependent change of K_{Mg} . K_{Mg} at different holding potentials (V) was obtained from experiments similar to that in *a*, and was plotted as a function of potential. Data were collected from a total of 25 cells. K_{Mg} increased as the membrane depolarized. The solid lines were drawn using the equation $K_{\text{Mg}} = K_{\text{Mg}0} \exp(zF\delta V/RT)$, in which $K_{\text{Mg}0}$ is the apparent dissociation constant at 0 mV membrane potential, z the valence, F the Faraday constant, δ the electrical distance, V the membrane potential, R the gas constant and T the temperature. Average $K_{\text{Mg}0} = 1,073.5 \text{ } \mu\text{M}$ ($a_{\text{K}_{\text{Mg}}} = 684.5 \text{ } \mu\text{M}$), $\delta = 0.73$ in control internal solution, $K_{\text{Mg}0} = 5,148.4 \text{ } \mu\text{M}$ ($a_{\text{K}_{\text{Mg}}} = 3,284.7 \text{ } \mu\text{M}$), $\delta = 0.81$ in the presence of PKC.

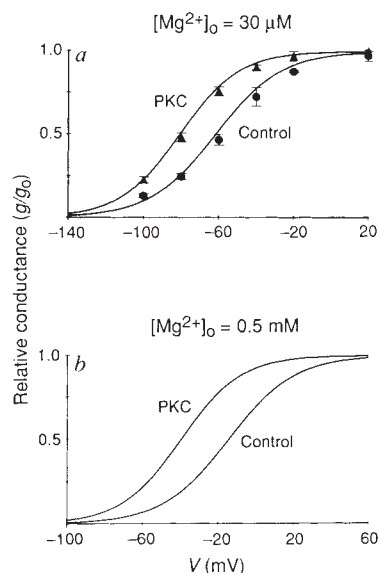


FIG. 4. Effect of PKC on conductance-voltage relationship of NMDA-receptor channels. *a*, The relative conductance g/g_0 of the NMDA-receptor channel in $30 \mu\text{M Mg}^{2+}$ was plotted as a function of holding potential (V). Data were from the same experiments as those shown in Fig. 3. Solid curves were drawn according to the equation $g/g_0 = 1 - [\text{Mg}^{2+}] / (K_{\text{Mg}^{2+}} \exp(zF\delta V/RT) + [\text{Mg}^{2+}])$. The parameters are as given in Fig. 3. The midpoint of the curve was -62.7 mV for the control and -81.2 mV for PKC. *b*, Theoretical curves, calculated using the same equation and parameter values, except that $[\text{Mg}^{2+}] = 0.5 \text{ mM}$. The midpoint of the curve was -14.3 mV for the control and -37.6 mV for PKC.

phosphorylation of target proteins to sustain cellular responses. In trigeminal neurons, the reduction of Mg^{2+} block by PKC is responsible for a sustained potentiation of NMDA responses by μ -opioid receptor agonist D-Ala²-MePhe⁴-gly-ol⁵-enkephalin (DAGO)⁴. The same modulatory mechanism may be important in inducing and/or maintaining long-term potentiation in other central neurons. □

Molecular cloning and expression of a rat V1a arginine vasopressin receptor

Alain Morel*, Anne-Marie O'Carroll,
Michael J. Brownstein & Stephen J. Lolait†

Laboratory of Cell Biology, National Institute of Mental Health, Building 36, Room 3A-17, Bethesda, Maryland 20892, USA

THE neurohypophyseal hormone arginine vasopressin has diverse actions¹⁻⁷, including the inhibition of diuresis, contraction of smooth muscle, stimulation of liver glycogenolysis and modulation of adrenocorticotrophic hormone release from the pituitary. Arginine vasopressin receptors are G protein-coupled and have been divided into at least three types⁸; the V1a (vascular/hepatic)^{1,9-11} and V1b (anterior pituitary)¹² receptors which act through phosphatidylinositol hydrolysis to mobilize intracellular Ca^{2+} , and the V2 (kidney) receptor^{1,13,14} which is coupled to adenylate cyclase. We report here the cloning of a complementary DNA encoding the hepatic V1a arginine vasopressin receptor. The liver cDNA encodes a protein with seven putative transmembrane domains, which binds arginine vasopressin and related compounds with affinities similar to the native rat V1a receptor. The messenger RNA corresponding to the cDNA is distributed in rat tissues known to contain V1a receptors.

Because *Xenopus* oocytes injected with rat liver messenger RNA exhibit a specific increase in intracellular Ca^{2+} in response to application of arginine vasopressin (AVP), this assay was used to isolate a functional cDNA clone encoding a V1a receptor. A rat liver cDNA library was made in a mammalian expression vector modified to permit RNA synthesis from the cDNA inserts (see Fig. 1 legend). Plasmid DNA from sublibraries was transcribed *in vitro* and the mRNA injected into albino frog oocytes. After 2-3 days, the Ca^{2+} -binding photoprotein aequorin was injected into the oocytes and they were tested for the presence of an AVP receptor by measuring light emission in response to application of AVP. Stepwise fractionations of the response-evoking cDNA pools identified a single recombinant clone, 15a-36, that confers an AVP-induced Ca^{2+} response in oocytes and AVP binding in transfected COS-7 cells (see Fig. 1 legend).

The 1,354-nucleotide sequence of the V1a AVP receptor cDNA encodes a 394-amino-acid protein (M_r 44,202) deduced from the longest open reading frame of the cloned cDNA (Fig. 1*a*). The nucleotide sequence surrounding the initiation codon fits the consensus initiation sequence¹⁵. The translated protein has seven clusters of 20-25 hydrophobic residues, predicted to represent membrane-spanning domains, connected by three extracellular and three intracellular loops. The structure is similar to bacterial rhodopsin and other G protein-coupled receptors. Amino-acid identities between the V1a receptor and other members of the G protein-coupled receptor superfamily are concentrated in the transmembrane domains (Fig. 1*b*). In these domains, the 15a-36 protein has a sequence identity of 24-30% with the rat D2 dopamine receptor¹⁶, human 5HT1a receptor¹⁷, rat substance P receptor¹⁸ and rat endothelin A receptor¹⁹. There are several possible sites for post-translational modification of the V1a receptor. The N-terminal region preceding the first putative transmembrane domain contains two potential N-linked glycosylation sites at Asn 14 and Asn 27. There are 12 and 8 threonine-serine residues in the third cytoplasmic loop and the C-terminal region of the V1a receptor, respectively. These residues could be sites for regulatory phosphorylation, suggesting that receptor function is regulated by protein kinases.

Received 23 December 1991; accepted 31 January 1992.

1. Dickenson, A. H. *Trends pharmac. Sci.* **11**, 307-309 (1990).
2. Woolf, C. J. & Thompson, S. W. N. *Pain* **44**, 293-299 (1991).
3. Dubner, R. & Bennett, G. J. A. *Rev. Neurosci.* **6**, 381-418 (1983).
4. Chen, L. & Huang, L.-Y. M. *Neuron* **7**, 319-326 (1991).
5. Yamamoto, D. J. *Physiol., Lond.* **400**, 691-700 (1988).
6. Gerber, G., Kangrga, I., Ryu, P. D., Larew, J. S. A. & Randic, M. *J. Neurosci.* **9**, 3606-3617 (1989).
7. Kelso, S. R., Nelson, T. E. & Leonard, J. P. *J. Physiol., Lond.* (in the press).
8. Hu, G.-Y. *et al. Nature* **328**, 426-429 (1987).
9. Malinow, R., Madison, D. V. & Tsien, R. W. *Nature* **335**, 820-824 (1988).
10. Muller, D., Joly, M. & Lynch, G. *Science* **242**, 1694-1697 (1988).
11. Linden, D. J. & Routtenberg, A. *Brain Res. Rev.* **14**, 279-296 (1989).
12. Malenka, R. C. *et al. Nature* **340**, 554-557 (1989).
13. Malinow, R., Schulman, H. & Tsien, R. W. *Science* **245**, 862-866 (1989).
14. Madison, D. V., Malenka, R. C. & Nicoll, R. A. *Rev. Neurosci.* **14**, 379-397 (1991).
15. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, Z. *Nature* **307**, 462-465 (1984).
16. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261-263 (1984).
17. Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* **394**, 501-527 (1987).
18. House, C. & Kemp, B. E. *Science* **238**, 1726-1728 (1987).
19. Rane, S. G., Walsh, M. P., McDonald, J. R. & Dunlap, K. *Neuron* **3**, 239-245 (1989).
20. Legendre, P. & Westbrook, G. L. *J. Physiol., Lond.* **429**, 429-449 (1990).
21. Jahr, C. E. & Stevens, C. F. *Nature* **325**, 522-525 (1987).
22. Cull-Candy, S. G. & Usowicz, M. M. *Nature* **325**, 525-528 (1987).
23. Ascher, P., Bregestovski, P. & Nowak, L. *J. Physiol., Lond.* **399**, 207-226 (1988).
24. Goldberg, M., Agus, C. & Goldfarb, S. in *Kidney* (eds Brenner, B. & Rector, F.) 344-390 (Saunders, Philadelphia, 1976).
25. Forsythe, I. D. & Westbrook, G. L. *J. Physiol., Lond.* **396**, 515-533 (1988).
26. Lester, R. A., Clements, J. D., Westbrook, G. L. & Jahr, C. E. *Nature* **346**, 565-567 (1990).
27. MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J. & Barker, J. L. *J. Neurosci.* **7**, 3230-3244 (1986).
28. Ascher, P. & Nowak, L. *J. Physiol., Lond.* **399**, 247-266 (1988).
29. Gu, Y. & Huang, L.-Y. M. *Neuron* **6**, 777-784 (1991).
30. Hamill, O., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch.* **391**, 85-100 (1981).
31. Fabiato, A. & Fabiato, F. *J. Physiol., Paris* **75**, 463-505 (1979).
32. Pitzer, K. S. & Kim, J. J. *J. Am. chem. Soc.* **96**, 5701-5707 (1974).
33. Bell, R. M., Hannun, Y. & Loomis, C. *Meth. Enzym.* **124**, 353-359 (1986).

ACKNOWLEDGEMENTS. We thank S. Y. Wong and K. Forni for technical assistance, and T. Fulford and M. Watson for preparing the manuscript. The work is supported by NIH grants to L.M.H. and a grant from Bristol Myers-Squibb Co.

* Present address: Group de Neurobiochimie Cellulaire et Moléculaire, 96 Bd Raspail, 75006 Paris, France.

† To whom correspondence should be addressed.

a

gggggggga -241
 atctaagactgcgtgggagctcagagagtgccccctgcctcaggaccagacagaagtaggacatgttgaatattccatttaaaccactctgcacgcacggattgcccacgct -121
 agagacggagggaggaaggaagagcagagcgcagagctcagaaactcggatcctccggtcctcggtggagaggaggaagcagccacggggcagggcctccagagctctgacggacagc -1

atgagtttcccgagagctccaggatcggtccgtgggaactccagcccggtggtgctctaacaccaggggctccaacggcagctcaggaggcagccaggtctgggaaggtgacagc 120
 MetSerPheProArgGlySerGlnAspArgSerValGlyAsnSerSerProTrpTrpProLeuThrThrGluGlySerAsnGlySerGlnGluAlaAlaArgLeuGlyGluGlyAspSer 40

-----I-----
 ccgctgggggagctacgcaatgaggagctggccaaactggaactcgctgtgctggcagctattttgtggtgctgctgggcaatagcagtgctgctggcgtgcatcgacgcca 240
 ProLeuGlyAspValArgAsnGluGluLeuAlaLysLeuGluIleAlaValLeuAlaValIlePheValValAlaValLeuGlyAsnSerSerValLeuLeuAlaLeuHisArgThrPro 80

-----II-----
 cgaagacatcccgatgcacactcttatccgacacctcagctggcagacctggcgctccttccaagtattaccgcagctatgctggacatcacctagcttccggggcgggac 360
 ArgLysThrSerArgMetHisLeuPheIleArgHisLeuSerLeuAlaAspLeuAlaValAlaPhePheGlnValLeuProGlnLeuCysTrpThrSerProSerPheArgGlyProAsp 120

-----III-----
 tggctgtgcgctggtgaagcactgcaggtgttgcctgttcgctctgcctatgctggtggtgatgacagccgacgctacatcgccgtgtgccacccgctcaagaccctgcag 480
 TrpLeuCysArgValValLysHisLeuGlnValPheAlaMetPheAlaSerAlaTyrMetLeuValValMetThrAlaAspArgTyrIleAlaValCysHisProLeuLysThrLeuGln 160

-----IV-----
 cagccggcgccgctcgccctcatgatcgccacctctgggtgctgagtttataactaacgacgacagctactttatattctctgtgatcgaaatcgaggtgaacaatggcactaaa 600
 GlnProAlaArgArgSerArgLeuMetIleAlaThrSerTrpValLeuSerPheIleLeuSerThrProGlnTyrPheIlePheSerValIleGluIleGluValAsnAsnGlyThrLys 200

-----V-----
 acccaagactgctgggctacettcatccagccctggggtaccgcgctacgtgacctggatgaccagcggtgtctctcggtggcactgtggtcgtcttgggtacatgctatggcttcac 720
 ThrGlnAspCysTrpAlaThrPheIleGlnProTrpGlyThrArgAlaTyrValThrTrpMetThrSerGlyValPheValAlaProValValValLeuGlyThrCysTyrGlyPheIle 240

-----VI-----
 tgctaccacatctggcgcaacatccgaggaaagacagcgtcctcgacacacagcaagggtgacaagggtctggggagggcggtgggtccctttcataaggggcttttggttacacctgt 840
 CysTyrHisIleTrpArgAsnIleArgGlyLysThrAlaSerSerArgHisSerLysGlyAspLysGlySerGlyGluAlaValGlyProPheHisLysGlyLeuLeuValThrProCys 280

-----VII-----
 gtcagcagcgtgaagagcatttcccgcccaagatccgcagctgtgaagatgacctttgtgattgtaagcgctacatccttctgctggcgccctttcttcatgtccagatgtggtcagtc 960
 ValSerSerValLysSerIleSerArgAlaLysIleArgThrValLysMetThrPheValIleValSerAlaTyrIleLeuCysTrpAlaProPhePheIleValGlnMetTrpSerVal 320

-----VIII-----
 tgggatgagaatttcatctggacgattcagaaaacccttccatcacaatcacggcgttgcgtcttctgaacagctgctgcaaccgctggatatacatgttttccagtgccatctc 1080
 TrpAspGluAsnPheIleTrpThrAspSerGluAsnProSerIleThrIleThrAlaLeuLeuAlaSerLeuAsnSerCysAsnProTrpIleTyrMetPhePheSerGlyHisLeu 360

-----IX-----
 ctgcaagactcgtccaaagtttccatgctgccacagcatggcgagaaattcccaagatgactccgatagcatgagccgaaagacagacttcttattctacaacgggagcccaac 1200
 LeuGlnAspCysValGlnSerPheProCysCysHisSerMetAlaGlnLysPheAlaLysAspSerAspSerMetSerArgLysThrAspPheLeuPhe*** 394

-----X-----
 gaacagcactgggatgtgaaggactcgccaaactctccaaatccatcagattcattcctgtctccacttgagccacgcttcacgcagccgactcttgggaggagacttttgtgtcca 1320
 gctagagatcgtgaagcagatgtgttatgtgg 1354

b

V1a M-----S F-----P R G S O D R S V G N S S F W W-----P L T T E G S-----N G S Q E
 D2 M-----P L M L S W Y D D D L R Q-----N W S R P F N G S E G
 5HT1a M-----V L S P G G N N T T S P P A F F E T Y G G-----N T T G
 SP M-----D N V L-----P M D-----S D L F P N I S T-----N T S E S
 ETA M E T F W L R L S F W V A L V G G V I S D N P E S Y S T N L S I H V D-----S V A T F H G T E L S F V V T T H Q P T N L A L P S N G S M H

-----I-----
 V1a A A R L G E G D S P L G D V R M E E L A K I E I A V L A V I F V V A V L G S S V L A L H R T P R K T S R M H L F I R H L S L A D L
 D2 K A D R P H Y-----N Y Y-----A M L L T L L I F I I V F G M V L V C H A V S R E K A-----L Q T T N Y L I V S L A V A
 5HT1a I S D-----V T V S Y Q V I-----T S L L L G T L I F C A V L G M A C V V A I A L E R S-----L Q N V A N Y L I G S L A V T
 SP N Q F V Q P T-----W Q I V L W A A A T Y I V I V T S V V G V V V I W I I L A H K R M-----R T V T N Y F L V N L A F A
 ETA N Y C P Q Q T K-----I T S-----A F K Y I N T V I S C T I F I V G M V G H A T L R I I Y Q N K C M-----R N G P N A L I A S L A L G

-----II-----
 V1a A V A F F Q-----V L F P Q L C-----N T S P S F R G P D-----W L C R V V K L Q V F A M F A S A Y M L V V M T A D R Y I A V C H F L K
 D2 D L L V A T L V M P V V V-----Y L E V V G E M K F S R I H C D I F V T L D V M M C T A S I L N L C A I S I D R Y T A V A M P M L
 5HT1a D L M V S V L V L F M A A-----L Y Q V L N K W T L G Q V T C D L F I A L D V L C C T S S I L H L C A I A L D R Y M A I T D F I D
 SP E A C M A A F N T V V N F-----T Y A V H N V W Y Y G L F Y C R F H N F F P I A A L F A S I Y S M T A V A F D R Y M A I I H P L
 ETA D L I Y V I D L P I N V F K L L A G R W P F E Q N D F G V P L C K L F P F L Q K S S V G I T V L N L C A L S V D R Y R A V A S W S R

-----III-----
 V1a T L Q Q-----P A R R S R L M I A T S W V L S F I L S T P Q Y F I F S V-----I E I E V N N G T K T Q D C W A T F I Q P W G T R A Y-----V
 D2 Y N T R Y S S K R R V T V M I A I V W V L S F T I S C P L L P G L N N-----T D Q N E C I I A N P A F V V Y S S
 5HT1a Y V N K R T-----P R P R A L I S L T W L I G F L I S I P M L G M R-----T P E D R S D P D A C T I S K D H G Y T I
 SP Q P R L S A T A T K V V I F V I W V L A L L L A F P Q G Y Y S T T-----E T M P S R V V C M I E W P E H P N R T Y E K A Y-----H I
 ETA V Q G I G I P L V T A I E I V S I W I L S F I L A I P E A I G F V M V P F E Y K G A Q H R T C M L N A T S K F M E F Y Q D V-----K D

-----IV-----
 V1a T M M T S G V F V A P V V V L Q T C Y G F I C Y H I W R N I R G T-----A X I R T V K M T F V I V S A Y I L C W A P F F I V Q M W
 D2 I V S F Y V P F I V T L L Y I K I Y I V L R K R R K R V N T-----K R-----K E K K A T Q M L A I V L G V F I I C W L P F F I T H I L
 5HT1a Y S T F-----G A F Y I P L L L M L V L Y G R I F R A A R F R I R-----K T-----R E R K T V R T L G I I M G T F I L C W L P F F I V A L V
 SP C V T V L I Y F-----L P L L V I G Y A Y T V V G I T L W A S E I P G D-----A K R K V V K M H I V V V C T F A I C W L P F H V F F L L
 ETA W W L F-----G Y F C M P L V C T A I F Y T L M T C E M L N R R N G S L-----Q R R E V A R T V I C L V V I F A L C W F P L H L S R I L

-----V-----
 V1a S-----V W D E N F I M-----T D-----S E N F-----S I T T A L L A S L N S C N P M I Y M F F-----S G H L L Q D C----- (30)
 D2 N-----I H C D C N I P F V L Y S A F T M L G Y V N S A V N P I I Y T T F N I E F R K A F M K I L H C
 5HT1a L-----P F C E S S C H M P T L L G A Y I N W L G Y S N S L L N P V I Y A Y F N K D F O N A F K K I I K C N F C R Q
 SP-----P Y I M P D-----L Y L K K F I Q Q V L A S M W L A M S S T M Y N P I I Y C C L N D R F R L G F K H A F R C C----- (84)
 ETA K K T V Y D E M-----D T N R C E L-----L-----S F L L L M D Y I G I N L A T M N S C I N P I A L Y F V S K K F K N C F Q S C L C C C----- (40)

The V1a receptor lacks an aspartate residue in the third transmembrane domain which is implicated in the binding of charged amine moieties in monoaminergic and muscarinic receptors, and a phenylalanine residue in the sixth hydrophobic domain, a characteristic that the receptor has in common with the thromboxane A₂ (ref. 20) and cannabinoid²¹ receptors. The V1a receptor is not closely related to any other G protein-coupled

receptor, except the oxytocin (see accompanying paper³¹) and V₂ AVP (S.J.L., A.-M.O'C., O. W. McBride, M. König, A.M. and M.J.B.; manuscript submitted) receptors.

To determine the binding characteristics of the V1a receptor, we used CHO cells stably transfected with clone 15a-36. The membranes of transfected CHO cells bound [³H]AVP in a saturable fashion with a dissociation constant, K_d , of 0.67 nM

FIG. 1 Sequence of the rat V1a AVP receptor and comparison with the sequences of other G protein-coupled receptors. *a*, Partial nucleotide sequence for the 15a-36 cDNA and deduced amino-acid sequence of the rat V1a AVP receptor. The initial stretch of guanine nucleotides represent the G tail produced during cDNA synthesis. The deduced amino-acid sequence is shown beneath the nucleotide sequence; amino-acid numbering begins with the first methionine of the long open reading frame. An in-frame stop codon in the 5' untranslated region is underlined. Putative *N*-linked glycosylation sites arrowheads. Positions of the putative transmembrane segments I-VII, dashed lines above the nucleotide sequence; the termini of each segment are tentatively assigned on the basis of the hydrophobicity profile and sequence comparison with other G protein-coupled receptors. The nucleotide sequence has been deposited in the EMBL DataBank (accession no. Z11690) *b*, Alignment of the amino-acid sequences of the rat V1a AVP receptor (V1a), rat D2 dopamine receptor (D2), human 5HT_{1a} receptor (5HT_{1a}), rat substance P receptor (SP) and rat endothelin receptor (ETA). Amino acids enclosed in solid lines and shaped represent residues that are identical in the V1a AVP receptor compared to the other receptors. Putative transmembrane segments are indicated by dashed lines labelled by roman numerals. Number of residues in the variable third cytoplasmic loop and at the C terminus are in parentheses.

METHODS. A 1.6×10^6 recombinant rat liver cDNA library was made in the pcDSP6/T7 plasmid cloning vector (M.J.B., unpublished), a derivative of the pcD vector of Okayama and Berg³⁰. Plasmid pcDSP6/T7 has an SV40 promoter and origin of replication, an SP6 phage promoter immediately upstream, and a T7 promoter, *Not*I, *Bst*XI, and *Xba*I sites downstream of the insert. These modifications allow the cDNA to be transcribed *in vitro*. The library was size-selected by linearizing with *Not*I; linearized plasmids

with inserts 1–4 kb long were used to construct a second library. The library was subdivided into 24 pools of ~50,000 clones and plasmid midpreps from each subdivision were linearized with *Not* I. Plasmid DNA (1–2 µg) from each pool was transcribed *in vitro* by SP6 polymerase (Promega). The mRNA synthesized was extracted with phenol-chloroform, ethanol-precipitated and dissolved in 5 µl distilled H₂O. The mRNA solution (50 nl, ~30–50 ng) was injected into defolliculated albino *Xenopus* oocytes. Following 48–72 h incubation at 18 °C, the oocytes were tested for the expression of the AVP receptor by mobilization of intracellular Ca²⁺ as detected by aequorin luminescence²⁵. The emission of photons after application of AVP on four oocytes was determined at 6 s⁻¹ intervals in a liquid scintillation counter. Of the 24 initial subdivisions, the one with the best response was subdivided into successive pools of 4,000, 300 and 50 clones. The 50 clones of the final subdivision were tested individually and two positive clones of identical size (~1.8 kb) were isolated. The response of oocytes injected with mRNA synthesized from these clones was concentration-dependent and detectable with as little as 10^{-10} M AVP, and completely blocked by the V1a antagonist, mVP (data not shown). Thirty-six clones of the final 50-clone pool, with insert size >1.5 kb, were individually screened by transient transfection²⁶ of ~600,000 COS-7 cells per 10 cm dish with 20 µg plasmid DNA from the respective clones, followed by binding on ~100 µg crude plasma membranes using [³H]AVP^{12,27–29}. Two clones that specifically bound the tritiated ligand (B_{max} 672 fmol per mg protein; K_d 1.1 nM) had the same size and restriction map to those obtained by the oocyte expression studies; both strands of the cDNA insert from one of these clones (15a-36) was sequenced by the dideoxynucleotide chain termination method using Sequenase (US Biochemical Corp.).

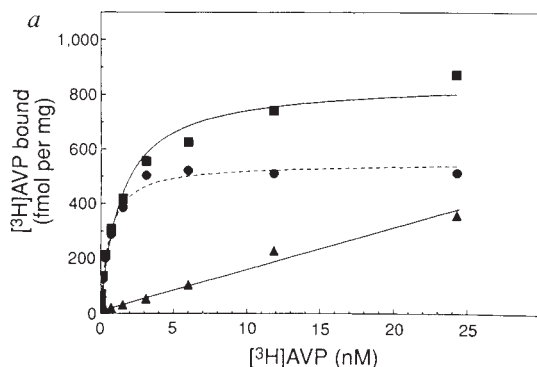
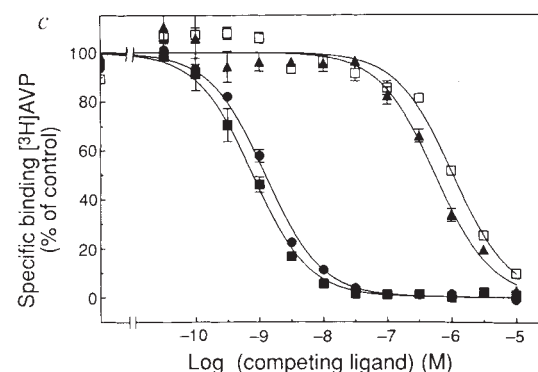
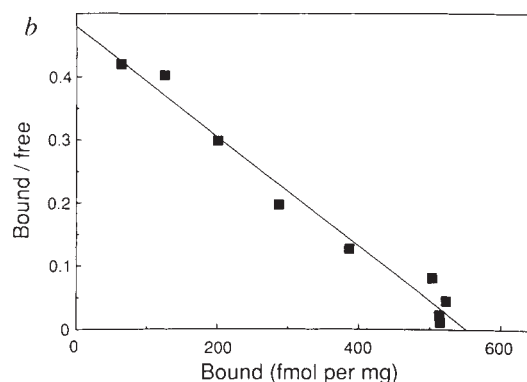


FIG. 2 Binding of [³H]AVP to transformed CHO-K1 cell membranes. *a*, Saturation isotherm of total (■), specific (●) and non-specific (▲) binding of [³H]AVP to membranes from CHO cells expressing the rat V1a receptor. Nonspecific binding was determined in the presence of 1 µM unlabelled AVP and was 8–16% of total binding when 0.5 nM [³H]AVP was used. *b*, The derived Scatchard plot. The estimated maximal binding, B_{max} , was 558 ± 13 fmol per mg protein (mean $\pm$ s.e.m., $n=3$) and the K_d 0.67 ± 0.06 nM. *c*, Competition of specific [³H]AVP binding to transformed CHO cell membranes. Representative curves are shown for the concentration-dependent inhibition of [³H]AVP binding by the indicated concentrations of AVP (●), mVP (■), oxytocin (▲) and dDAVP (□). Peptides were from Peninsula. Competition experiments used 0.5 nM [³H]AVP.

METHODS. The V1a receptor cDNA (15a-36) was introduced into CHO-K1 cells by calcium phosphate-mediated transfection³⁰. A clonal cell line expressing the 15a-36 cDNA was obtained by limiting-dilution cloning of cells expressing the corresponding mRNA, determined by northern blot analysis (Fig. 3). Confluent cells, collected by scraping, were homogenized in ice-cold 50 mM Tris-HCl, pH 7.5, containing 1 mM EGTA and 3 mM MgSO₄. A crude membrane fraction was prepared by centrifugation at 39,000g for 30 min at 4 °C; the pellet was rehomogenized and stored frozen in liquid N₂ at 1–2 mg ml⁻¹ protein concentration in the homogenizing buffer containing 0.1% BSA and 100 µg ml⁻¹ bacitracin. Binding of [³H]AVP to 50–100 µg of these membranes was done as described previously^{12,27–29} with the exception that the incubation volume was 200 µl and the GF/C filters were counted



in 10 ml Econofluor 2 (NEN) and a LKB RackBeta liquid scintillation counter with 58.8% efficiency. Results shown are the means of triplicate determinations and are representative of three experiments. Both saturation and competition binding data were analysed by a nonlinear least-squares curve-fitting program according to the method of Marquardt (InPlot, GraphPad, San Diego, USA).

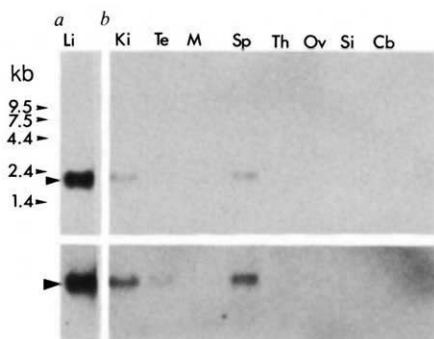


FIG. 3 Northern blot analysis of mRNA from rat tissues with V1a receptor oligodeoxynucleotides as probes. The ~2.1 kb V1a receptor mRNA is indicated by an unlabelled arrowhead. RNA size markers are shown, left (kb). Each lane contains 5 μ g poly(A)⁺ RNA from a, rat liver (Li), and b, rat kidney (Ki), testis (Te), spleen (Sp), thymus (T), ovary (Ov), small intestine (Si) and cerebellum (Cb). An overexposure of the autoradiograph is shown in the lower portion of the figure, emphasizing the presence of V1a transcript in the testis.

METHODS. Total RNAs were isolated from adult (150–180 g) S/D rat tissues using guanidinium thiocyanate³⁰ and poly (A)⁺ RNA isolated using oligo-(dT) cellulose columns. Poly (A)⁺ RNA (5 μ g of each) was electrophoresed through 6.6% formaldehyde/1% agarose, transferred to a Nytran nylon membrane (Schleicher & Schuell) and cross linked by ultraviolet light. The filter was prehybridized at 37 °C overnight in 4 $\times$ SSPE, 50% formamide, 5 $\times$ Denhardt's solution, 500 μ g ml⁻¹ sheared salmon sperm DNA, 250 μ g ml⁻¹ yeast transfer RNA and 0.1% SDS, and hybridized in the same buffer at 37 °C for 18 h with two 48-base oligodeoxynucleotide probes labelled on the 3'-end with terminal deoxynucleotidyl transferase and deoxyadenosine-[³²P]-triphosphate (NEN, >3,000 Ci mmol⁻¹). The probes are complementary to sequences encoding amino acids 1–16 and 271–286 of the rat V1a receptor, and correspond to regions in the N terminus and third cytoplasmic loop, respectively. These segments have the lowest degree of homology among different classes and subtypes of G protein-coupled receptors. After hybridization, the blots were washed at 56 °C in 1 $\times$ SSC/0.1% SDS before being exposed for 11 days to X-ray film at -80 °C with intensifying screens. This experiment was done twice and gave similar results.

and expressed quantities of receptor protein ranging from 540 to 584 fmol per mg protein (Fig. 2a, b). No such binding was detected on membranes prepared from untransfected cells or from cells transfected with the vector DNA alone (data not shown). Binding specificity was analysed by competition studies between [³H]AVP and other hormones and analogues with differing selectivities for the neurohypophyseal hormone receptors (Fig. 2c). The V1a pressor antagonist mVP(1-mercapto- β , β -cyclopentamethylene propionic acid (O-methy-tyrosine) 8-arginine) vasopressin, oxytocin and the V2 antidiuretic agonist dDAVP (1-deamino, 8-D-arginine vasopressin) competed for the [³H]AVP binding to transfected CHO membranes with an order of potency similar to that observed with the V1a receptor in rat liver membranes¹². The most potent inhibitor was mVP, being ~500-fold and >1,000 times more potent than oxytocin and dDAVP, respectively. The values of the inhibition constant (K_i) for mVP, AVP, oxytocin and dDAVP, were calculated to be 0.94 ± 0.08 , 1.82 ± 0.54 , 504.7 ± 58 and $1,143 \pm 43$ nM, respectively. Both the K_d and K_i values obtained for the cloned V1a AVP receptor agree with those for rat liver V1a receptors¹².

Northern blots of poly(A)⁺ mRNA from rat tissues were hybridized with synthetic oligodeoxynucleotide probes to the 15a-36 insert. A ~2.1-kilobase (kb) mRNA transcript was detected in the liver, kidney and spleen with smaller amounts in the testis; no detectable RNA species were observed in the thymus, ovary, small intestine or cerebellum (Fig. 3). *In situ* hybridization histochemistry revealed the presence of V1a receptor mRNA in liver, kidney and brain (data not shown). Specific V1a-binding sites and/or pharmacological effects have been demonstrated in all the tissues exhibiting the ~2.1 kb hybridizing band^{22–24}.

When expressed in *Xenopus* oocytes or eukaryotic cells, the 15a-36 cDNA encodes a V1a AVP receptor with the primary structure of a G protein-coupled receptor. Many other G protein-coupled receptors consist of multiple receptor subtypes which can be distinguished by their ligand selectivity, the second messengers involved in mediating their biological responses and their tissue distribution. Screening of rat and human genomic DNA digested with *Hind*III or *Eco*RI and hybridized at moderate stringency (3 $\times$ standard saline citrate, 60 °C) with a probe corresponding to transmembrane regions I to V of the rat V1a receptor cDNA revealed 5–9 bands in each digest (data not shown), indicating the existence of additional members of the vasopressin receptor gene family. Cloning of the V1a AVP receptor cDNA could therefore facilitate the isolation of other AVP isoreceptors. □

Received 6 November 1991; accepted 19 February 1992.

- Jard, S. *Front. Hormon. Res.* **13**, 89–104 (1985).
- van Ree, J. M., Bohus, B., Versteeg, D. H. G. & de Wied, D. *Biochem. Pharmac.* **27**, 1793–1800 (1978).
- Gillies, G., Linton, E. A. & Lowry, P. J. *Nature* **299**, 355–357 (1982).
- de Wied, D. *Life Sci.* **20**, 192–294 (1977).
- Kasting, N. W., Veal, W. C. & Couper, K. E. *Can. J. Physiol. Pharmac.* **58**, 316–319 (1980).
- Matsuguchi, H., Schmid, P. G., Gorden, F. J. & Johnson, A. K. *Physiology* **23**, 29–33 (1980).
- Clements, J. A. & Funder, J. W. *Endocr. Rev.* **116**, 416–423 (1986).
- Michell, R. H., Kirk, C. J. & Billah, M. M. *Biochem. Soc. Trans.* **7**, 861–865 (1979).
- Kirk, C. J., Verrinder, T. R. & Hems, D. A. *FEBS Lett.* **83**, 267–271 (1977).
- Aiyar, N., Nambi, P., Stassen, F. L. & Crooke, S. J. *Life Sci.* **39**, 37–45 (1986).
- Nambi, P., Whitman, M., Gessner, G., Aiyar, N. & Crooke, S. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8492–8495 (1986).
- Antoni, F. A. *Neuroendocrinology* **39**, 186–188 (1984).
- Chase, L. R. & Aurbach, G. D. *Science* **159**, 545–547 (1968).
- Butten, D. et al. *Molec. Pharmac.* **14**, 1006–1017 (1978).
- Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
- Bunzow, J. R. et al. *Nature* **336**, 783–787 (1988).
- Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
- Bunzow, J. R. et al. *Nature* **336**, 783–787 (1988).
- Kobilka, B. K. et al. *Nature* **329**, 75–79 (1987).
- Yokota, Y. et al. *J. biol. Chem.* **264**, 17649–17652 (1989).
- Arai, H., Hori, S., Aramori, I., Ohkubo, H. & Nakanishi, S. *Nature* **348**, 730–732 (1990).
- Hirata, M. et al. *Nature* **349**, 617–620 (1991).
- Matsuda, L. A., Lolait, S. J., Brownstein, M. J., Young, A. C. & Bonner, T. I. *Nature* **346**, 561–564 (1990).
- Phillips, P. et al. *Endocrinology* **126**, 1478–1484 (1990).
- Meidan, R. & Hsueh, A. J. W. *Endocrinology* **116**, 416–423 (1985).
- Thibonnier, M., Snajdar, R. M. & Rapp, J. P. *Am. J. Physiol.* **251**, 115–120 (1986).
- Woods, N. M., Cuthbertson, K. S. R. & Cobbold, P. H. *Nature* **319**, 600–602 (1986).
- Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
- Holmes, M. C., Antoni, F. A. & Szentendrei, T. *Neuroendocrinology* **39**, 162–169 (1984).
- Antoni, F. A. *Neuropeptides* **4**, 413–420 (1984).
- Balla, T., Enyedi, P., Spat, A. & Antoni, F. *Endocrinology* **117**, 421–423 (1985).
- Okayama, H. et al. *Meth. Enzym.* **154**, 3–28 (1987).
- Kimura, T., Tanizawa, O., Mori, K., Brownstein, M. J. & Okayama, H. *Nature* **356**, 526–529 (1992).

ACKNOWLEDGEMENTS. We thank L. Mahan for assistance in analysing the binding data. A.M. was a Visiting Fellow of the Fogarty International Centre, A.-M. O.C. was supported in part by a grant from Ciba Geigy, and S.J.L. by grants from the National Alliance for Research on Schizophrenia and Depression and the Stanley Foundation.

Structure and expression of a human oxytocin receptor

Tadashi Kimura*, Osamu Tanizawa*, Kensaku Mori†, Michael J. Brownstein‡ & Hiroto Okayama§

* Department of Obstetrics and Gynecology, Osaka University Medical School, 1-1-50 Fukushima, Fukushima-ku, Osaka 553, Japan

† Department of Neuroscience, Osaka Bioscience Institute, Suita, Osaka 565, Japan

‡ Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20892, USA

§ Department of Molecular Genetics, Osaka University Research Institute for Microbial Diseases, Suita, Osaka 565, Japan

JUST before the onset of labour, uterine myometrium becomes extremely sensitive to oxytocin¹, for which it is a primary target tissue, because of a dramatic increase in the number of oxytocin receptors^{2,3}. We report here the structure and expression of the human oxytocin receptor complementary DNA isolated by expression cloning. The encoded receptor is a 388-amino-acid

polypeptide with 7 transmembrane domains typical of G protein-coupled receptors. The oxytocin receptor, expressed in *Xenopus* oocytes, specifically responds to oxytocin and induces an inward membrane current. Messenger RNAs for the receptor are of two sizes, 3.6 kilobases in breast, and 4.4 kilobases in ovary, uterine endometrium and myometrium. The mRNA level in the myometrium is very high at term. We conclude that the increase in receptor number in the myometrium at labour is, at least in part, due to the increase in mRNA.

Some G protein-coupled membrane receptors elicit a ligand-specific electrophysiological response following their expression in *Xenopus* oocytes⁴. When we injected defolliculated *Xenopus* oocytes with mRNA prepared from myometrium, we found that oxytocin produced a large inward membrane current (Fig. 1a) followed by oscillation⁵ as has been reported⁶. Using this phenomenon as an assay of size-fractionated mRNA⁷, we determined the receptor mRNA to be 4–5 kilobase (kb) in size (data not shown). To clone the cDNA for the receptor, we constructed a cDNA library with total myometrium mRNA using a vector containing different RNA polymerase promoters upstream and downstream of the cDNA inserts. A size-fractionated sub-library

containing cDNA inserts greater than 4 kb was then prepared and divided into 16 pools of approximately 6,000 cDNA clones. Each pool was transcribed *in vitro* and the RNA products were injected into *Xenopus* oocytes (Fig. 1a)^{8–10}. We found that RNA from one pool rendered oocytes responsive to oxytocin. After screening progressively smaller sub-libraries containing 1,500, 150, and 16 clones, we finally isolated a single clone (pOTR)

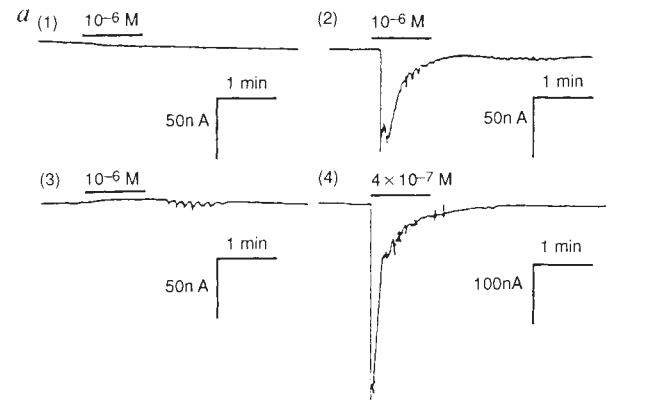


FIG. 1 a, Oxytocin-induced membrane currents recorded from voltage clamped (−60 mV) *Xenopus* oocytes injected with (1) water, (2) 50 ng poly(A)⁺ myometrial RNA, (3) 25 ng RNA transcribed *in vitro* from a positive pool of 6,000 clones, (4) 2 ng *in vitro* transcript of pOTR. The indicated oxytocin concentrations were applied to the oocyte bath for the time shown by bars. b, The primary structure of the human oxytocin receptor cDNA and encoded protein. The deduced amino acid sequence is shown above the cDNA sequence. Lines are above seven putative transmembrane domains (I–VII); the terminus of each domain is tentatively assigned according to the hydrophobicity profile and by comparison with other G protein-coupled receptors. Postulated N-linked glycosylation sites (◆) and putative phosphorylation sites (▼) are also shown. The dotted line indicates residual undetermined sequences (~50 bp) before the poly(A) tail.

METHODS. Human myometrium at term was obtained from a surgical specimen of a ruptured uterus with informed consent of the patient. Total RNA was prepared by the guanidine isothiocyanate/cesium trifluoroacetate method²⁰. Poly(A)⁺ RNA was purified using oligo(dT)-cellulose. A full length cDNA library of 10⁷ independent clones was constructed by the Okayama-Berg method²⁰ with the pCDSP6/77 vector. This plasmid is derived from the original pCD expression vector, which it differs in having an SP6 RNA polymerase promoter upstream of the cDNA and a T7 polymerase promoter downstream. In addition, there are three unique rare restriction sites (*not* I, *Bst* XI and *Xba* I) downstream of the insert in pCDSP6/77. A sub-library with inserts larger than 4 kb was constructed by digestion with *Not*I followed by size fraction-

ation on a 1% low melting agarose gel, recircularization and re-transfection into *E. coli* (DH-5). Plasmid DNA (5 µg) was prepared from each pool of 6,000 clones and digested with *Not*I or *Bst*XI. Digested DNA was then transcribed with SP6 RNA polymerase (Takara) in the presence of a cap analogue 7m-GpppG as previously described²¹. Electrophysiological responses to oxytocin (10^{−6} M, Sigma) were measured following injection of *in vitro* synthesized mRNAs into oocytes. Oocytes were defolliculated by collagenase treatment as described². The nucleotide sequence of pOTR was determined by sequencing both strands as described previously²².

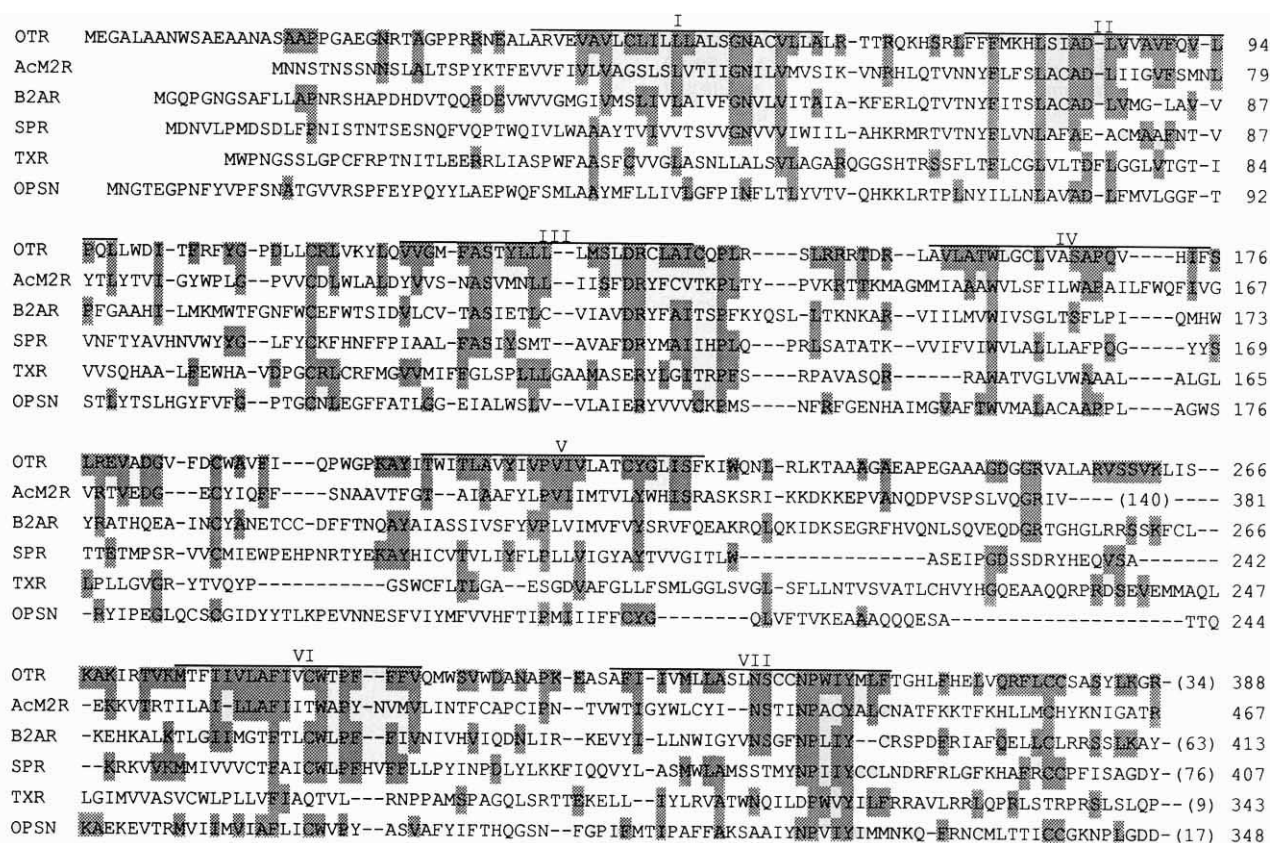


FIG. 2 Alignment of the amino acid sequences of seven-transmembrane receptor proteins. Deduced amino-acid sequences of the human oxytocin receptor (OTR), human muscarinic M2 receptor²³ (AcM2R), human β adrenergic receptor²⁴ (B2AR), rat substance P receptor²⁵ (SPR), human thromboxane

A2 receptor²⁶ (TXR) and human rhodopsin²⁷ (OPSN) were aligned. The putative transmembrane domains I-VII are indicated by lines above the sequences and conserved amino acids are shaded.

which seemed to encode the human oxytocin receptor (Fig. 1a).

The clone pOTR contains a 4,100 base-pair (bp) cDNA (Fig. 1b). The putative initiation codon is preceded by an in-frame termination codon and the surrounding nucleotide sequence is in good agreement with the Kozak consensus sequence¹¹. We therefore concluded that the cDNA contained the entire coding region. The cDNA encodes a 388-amino-acid protein with a relative molecular mass (M_r) of 42,716. The hydropathy profile¹² of the deduced amino-acid sequence suggests that the protein has 7 putative transmembrane domains, 20–30% identical to the corresponding domains of other G protein-coupled rhodopsin-type receptors (Fig. 2). In addition, the protein contains several sites for possible post-translational modifications. There are three consensus *N*-glycosylation sites¹³ in the amino-terminal region preceding the first putative transmembrane domain and possible phosphorylation sites¹⁴ are present in the third cytoplasmic loop and in the carboxy-terminal region (Fig. 1b).

To confirm the ligand specificity of the cloned receptor, we investigated the electrophysiological responses to its expression in *Xenopus* oocytes. Both oxytocin and Arg⁸-vasopressin (AVP) elicited inward currents with half-maximal concentrations of $\sim 10^{-8}$ M and 10^{-6} M, respectively (Fig. 3). This was not unexpected; the oxytocin receptor is known to respond weakly to vasopressin. A vasopressin V_1 receptor agonist¹⁵, [Phe², Ile³, Orn⁸]-vasopressin, had no effect at concentrations as high as 10^{-5} M, however (Fig. 3). When 10^{-6} M [d(CH₂)₅Tyr(OMe)²Orn⁸]-vasotocin¹⁶, a specific oxytocin antagonist, was added to the perfusate 1 min before oxytocin was applied to the oocyte, the amplitudes of the membrane currents were markedly decreased (Fig. 3). Thus, the cloned receptor is very similar in ligand specificity to the oxytocin receptor in uterine myometrium.

Northern blot analysis revealed that myometrium at term

contained an enormous amount of oxytocin receptor mRNA whereas early gestational or non-pregnant myometrium had little or none (Fig. 4). In term myometrium, 4 hybridizing bands with estimated sizes of 6.1, 4.4, 3.6 and 2.6 kb were detected. These mRNAs are likely to be derived from the same receptor gene because all the bands were visualized when 2 hybridization probes corresponding to the 5' and 3' terminal regions of the coding sequence were used for blot analysis at high stringency (data not shown). It is possible that the 6.1 kb and 2.6 kb mRNAs are, respectively, immature unspliced mRNA and degraded mRNA since these mRNAs were not detected in other tissues even after a long exposure (3 weeks). The cloned cDNA corresponds to the 4.4 kb mRNA that is most abundant in term myometrium (Fig. 4).

Oxytocin receptors are expressed in non-pregnant endometrium, myometrium and mammary gland but their numbers vary with the stage of the menstrual cycle¹⁷. The endometrial sample we analysed contained the 4.4 kb mRNA but the myometrial one did not (Fig. 4). This discrepancy may be a reflection of different menstrual stages of the patients from which the samples were obtained. Interestingly, the mammary gland expressed the 3.6 kb mRNA that corresponds to the third largest hybridizing band in term myometrium (Fig. 4). Whether this mRNA encodes the same oxytocin receptor as the 4.4 kb species remains to be determined. The presence of oxytocin in non-pregnant ovary has been reported but the expression of the receptor in this organ is controversial^{18,19}. We detected receptor mRNA in the ovary of a non-pregnant patient (Fig. 4). This clearly indicates that ovary expresses the oxytocin receptor gene but whether functional oxytocin receptor proteins are produced remains to be shown. Oxytocin receptor mRNA was not detected in cerebral cortex, liver, kidney and spleen.

We conclude that we have indeed cloned the oxytocin receptor

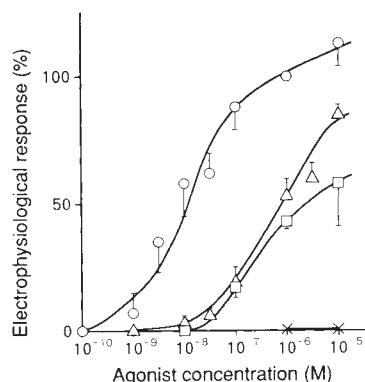


FIG. 3 Dose-response curves to oxytocin (○), Arg⁸-vasopressin (Δ), [Phe², Ile³, Orn⁸]-vasopressin, (×) and oxytocin (□) after pretreatment for 1 min with 10⁻⁶ M [d(CH₂)₅-Tyr(OMe)²Orn⁸]-vasotocin oxytocin-specific antagonist.

METHODS. *In vitro* transcribed pOTR (2 ng) was injected into defolliculated oocytes which were then incubated for 1–2 days at 19 °C. Responses to the ligands were recorded under voltage clamp at -60 mV. Vasopressins and vasotocin were from Peninsula Laboratories. After each ligand was applied and electrophysiological responses recorded, oocytes were washed for 15 min and 10⁻⁶ M oxytocin was applied as a control, as 100% in each case. The mean amplitude of membrane currents induced by 10⁻⁶ M oxytocin was 776 ± 44.0 nA (mean ± s.e.m. *n* = 60; range: 250–1350 nA). Each point represents the mean of at least three sets of data; bars represent s.e.m.

and that the marked increase in the number of oxytocin receptors in myometrium at the time of labour is, at least in part, a result of a marked elevation of mRNA level. Since the number of receptors seems to play a major role in the onset of labour, it will be interesting to know how the expression of the oxytocin receptor gene is regulated at parturition. The cloned cDNA should allow us to address this question and provides an opportunity to understand the physiological role of the oxytocin receptor that is expressed in the ovary.

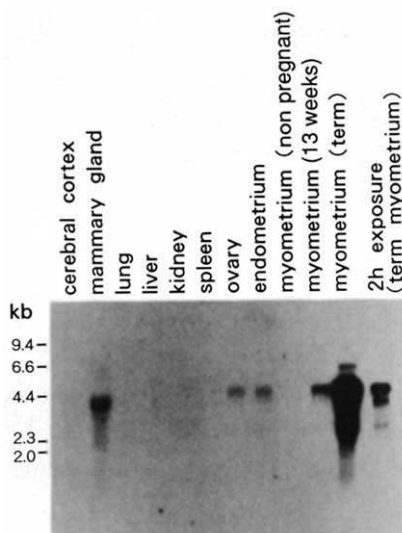


FIG. 4 Northern blot analysis of RNA isolated from various human tissues. Lung, liver and kidney were obtained at autopsy. Other tissues were from surgical specimens obtained with informed consent. The term myometrium sample used was described in Fig. 1.

METHODS. Total RNA (10 μg; cerebral cortex and mammary gland) or 2 μg poly(A)⁺ RNA (other tissues) were electrophoresed on a 1% agarose–2.2 M formaldehyde gel. After blotting to a nylon membrane filter (0.22 μm; Biohyne, Pall), the filter was prehybridized for 4 h and hybridized at 42 °C for 16 h with [³²P]-dCTP-labelled 0.8 kb BamHI–PstI fragment of the cDNA. The filter was washed for 1 h at 50 °C in 0.2 × SSC buffer, 0.1% SDS and exposed to a Kodak XAR film for 16 h or 2 h with 2 intensifying screens at -70 °C.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X64878. □

Received 18 November 1991; accepted 14 February 1992.

1. Cunningham F. G. MacDonald, P. C. & Gant, N. F. in *Williams Obstetrics* 18th (edn.) 187–226 (Appleton & Lange, Norwalk, Connecticut/San Mateo, California, 1989).
2. Soloff, M. S., Alexandrova, M. & Fernstrom, M. A. *Science* **204**, 1313–1315 (1979).
3. Fuchs, A. R., Fuchs, F., Husslein, P. & Soloff, M. S. *Am. J. Obstet. Gynecol.* **150**, 734–741 (1984).
4. Sigel, E. *J. Membrane Biol.* **117**, 201–221 (1990).
5. Kimura, T. et al. *J. Steroid Biochem. molec. Biol.* (in the press).
6. Morley, S. D., Meyerhof, W., Schwarz, J. & Richter, D. *J. molec. Endocr.* **1**, 77–81 (1988).
7. Hirano, C., Ito, I. & Sugiyama, H. *J. Physiol., Lond.* **382**, 523–535 (1987).
8. Masu, Y. et al. *Nature* **329**, 836–838 (1987).
9. Julius, D., MacDermott, A. B., Axel, R. & Jessell, T. M. *Science* **241**, 558–564 (1988).
10. Masu, M., Tanabe, Y., Tsuchida, K., Shigemoto, R. & Nakanishi, S. *Nature* **349**, 760–765 (1991).
11. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8148 (1987).
12. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
13. Hubbard, S. C. & Ivatt, R. J. A. *Rev. Biochem.* **50**, 555–583 (1981).
14. Kennelly, P. J. & Krebs, E. G. *J. biol. Chem.* **266**, 15555–15558 (1991).
15. Huguénin, R. L. *Helv. chim. Acta* **47**, 1934–1941 (1964).
16. Bankowski, K., Manning, M., Seto, J. & Sawyer, W. H. *Int. J. Peptide Protein Res.* **16**, 382–391 (1980).
17. Fuchs, A. R. et al. *Endocrinology* **127**, 629–636 (1990).
18. Sernia, C., Gemmell, R. T. & Thomas, W. G. *J. Endocr.* **121**, 117–123 (1989).
19. Fuchs, A. R. et al. *Am. J. Obstet. Gynecol.* **163**, 1961–1967 (1990).
20. Okayama, H. et al. *Meth. Enzym.* **154**, 3–28 (1987).
21. Noma, Y. et al. *Nature* **319**, 640–646 (1986).
22. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
23. Peralta, E. G. et al. *EMBO J.* **6**, 3923–3929 (1987).
24. Kobilka, B. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 46–50 (1987).
25. Yokota, Y. et al. *J. biol. Chem.* **264**, 17649–17652 (1989).
26. Hirata, M. et al. *Nature* **349**, 617–620 (1991).
27. Nathans, J. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4851–4855 (1984).

ACKNOWLEDGEMENTS. We thank T. Inoue, M. Takemura and H. Inoue for technical assistance.

Cloning and characterization of a gene that regulates cell adhesion

W. E. Pullman & W. F. Bodmer

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

MOLECULES of the cadherin and integrin families involved in cell–cell and cell–matrix adhesion have been implicated in epithelial differentiation, carcinogenesis and metastasis^{1–6}. Having observed that a colon cancer cell line bound avidly to collagen type I, inducing integrin-triggered glandular differentiation⁷, we investigated the regulation of integrin function in these cells. We modified a mammalian expression cloning system⁸ that used monoclonal antibody selection to clone cell surface molecules. Using attachment to collagen type I to select for adhesive phenotype, we isolated a complementary DNA clone that increases cell adhesion to components of the extracellular matrix. The corresponding gene (cell adhesion regulator, *CAR*) is located on the long arm of chromosome 16 (16q) and encodes a protein of 142 amino acids, which has an N-terminal myristoylation motif and a consensus tyrosine-kinase phosphorylation site at the C terminus. Removal of this tyrosine residue abolishes enhancement of cell–matrix adhesion. This gene may encode an adhesion signal transduction molecule that functions in the suppression of tumour invasion.

Our cloning strategy used WOP cells, which support extra-chromosomal replication of the pCDM8 vector⁸, because they have only low levels of binding to collagen type I (about 10%). An SW1222 colon cancer cell line cDNA library in pCDM8 was enriched for cDNAs that significantly enhanced the attachment of WOP cells to collagen type I after four rounds of selection (Fig. 1a). In particular, cDNA clone 27 induced three- to five-fold higher levels of adhesion, and the induced adhesion was specifically inhibited by the RGD (Arg-Gly-Asp) integrin-recognition sequence (Fig. 1b; mean 73% adhesion inhibition).

This approach could be used with other parameters to isolate novel adhesion molecules or their regulators, for which monoclonal antibodies are not available. The major limitation is the selection of the appropriate cell type with which to transfect the cDNA; it must allow high-copy-number plasmid replication

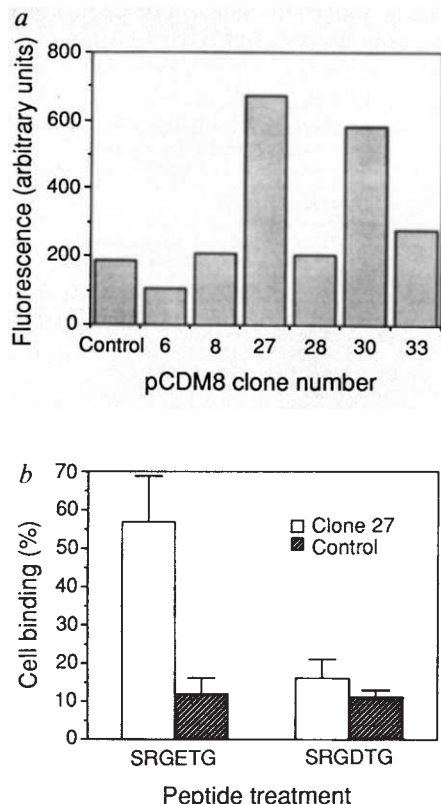


FIG. 1 *CAR* induces cell adhesion to collagen type I. *a*, Collagen type I adhesion assay⁷ used for initial screening of six independent cDNAs, isolated from an SW1222 library, which induce collagen binding in WOP cells ($n=2$ independent experiments; binding is proportional to fluorescence). *b*, WOP cells transfected with *CAR* (clone 27; open bars) have 4–5-fold enhanced collagen type I adhesion, compared with control cells (CD2-transfected; hatched bars) and binding is inhibited by RGD-containing peptide. Results in *b* are expressed as the percentage of cells that bind; the data are mean $\pm$ s.e.m., from 6 independent experiments.

METHODS. A cDNA library was made from mRNA of 5×10^7 SW1222 cells using standard methods²³. After ligation into the pCDM8 vector⁸ and expression in *Escherichia coli*, an aliquot of the resulting caesium chloride-purified plasmid library was electroporated into WOP cells (25 μ l of plasmid DNA into 10^7 cells at 300 V). The efficiency of transfection was 30–40%, judged by control transfection of WOP with CD2 in vector pCDM8. Expression of CD2 was detected by fluorescence microscopy with anti-CD2 (Becton-Dickinson) and FITC-conjugated F(ab)₂ antibodies (Dakopatts). After 48 h growth, WOP cells were selected for adhesion to collagen type I-coated dishes, prepared overnight with collagen type I solution (20 μ g ml⁻¹ in PBS) under sterile lamina flow at room temperature. Dishes were washed three times with PBS before use. Freshly trypsinized WOP cells (1×10^6) were washed twice in DMEM with BSA (2 mg ml⁻¹), applied to a 10-cm dish and incubated for 2 h at room temperature. Non-adherent cells were washed off by streaming PBS over the plate three times and remaining adherent cells lysed *in situ* with SDS-EDTA to form a supernatant⁸. After recovery of plasmids and amplification in *E. coli*, the selection process was repeated three times before individual clones were selected and transfected into WOP cells for collagen binding assay as previously described⁷. The effect of blocking the RGD recognition sequence was tested by co-incubation of cells with synthetic peptides: SRGDTG or control SRGETG (single-letter amino-acid code), at 1 mg ml⁻¹ in the adhesion assay. Plates were washed gently 3 times with PBS and residual cell numbers determined by fluorescence assay using 4-methyl-umbelliferyl phosphate to detect cellular phosphatase activity²⁴. After subtraction of non-specific background adhesion to BSA-coated control wells, results were initially expressed as arbitrary fluorescence units (*a*) but subsequently, to allow direct comparisons between experiments, as a percentage of the input number of cells (maximum fluorescence) (*b*).

and permit selection in an adhesion assay by virtue of acceptably low levels of constitutive binding, so that a specific effect can be distinguished from background.

The sequence of clone 27 cDNA has an open reading frame of 426 bases, starting from a putative ATG transcription initiation site, predicting a protein of 142 amino acids (Fig. 2) encoded by the gene we call *CAR*. On searching the PIR 27, Swiss Protein and OWL databases, no similarity was found to sequences of any of the known α and β integrin chains. A protein motif search of the PROMOT database⁹ revealed a very good match for a tyrosine kinase phosphorylation site¹⁰ at the terminal tyrosine residue (Tyr 142), located immediately before the stop codons. The N terminus of the protein *CAR* has a myristoylation consensus signature¹¹, suggesting a cytoplasmic sub-membrane location for the protein, particularly as no trans-membrane region or leader sequence has been identified. *CAR* is thus unlikely to be a conventional adhesion molecule.

The significance of the terminal tyrosine as a possible phosphorylation site involved in control of *CAR* function was investigated by site-directed mutagenesis of clone 27, converting Tyr 142 to an additional stop codon. Transfection of this truncated construct into WOP cells did not lead to enhanced collagen-binding activity, in contrast to the unmodified clone, suggesting that the functional activity of the *CAR* product depends on phosphorylation of its terminal tyrosine (Fig. 3). This suggestion is consistent with the functional activation by tyrosine phosphorylation of p60^{src} (ref. 12) and of the fibronectin receptor¹³.

Southern blot analysis of human genomic DNA shows that the *CAR* sequence is located within two *Eco*R1 fragments of 4 and 3.5 kilobases with cross-reaction to mouse genomic DNA, as expected for an evolutionarily conserved sequence. A panel of human $\times$ rodent hybrid DNAs¹⁴ was used to map *CAR* to human chromosome 16q by Southern blotting (data not shown). This is not the location of known integrin chains: α 1 and α 2

```

ATGGGGAAGGTTATCAGTGCTTCCGAGTGAGCATGGAACACTTCGAG      48
 M G K V I S A S R V S M E H F E                               16

TTCCAGGGTTATAGACAGTCGTTCCAGTGCGGTGAGGCCACCCAGA      96
 F Q G Y R Q S F P V W L R P P R                               32

GGCAGCAGAGCATTGAGCTCCAAACAGACCCCTGTTTCATGCCGATGC      144
 G S R A F R L Q T D P C S C R C                               48

TTGCACGACCGCCCCAGTTCCTGTGGCTCCCTCGGAATGCTAAGGGGA      192
 L H D R P S S C G S L G M L R G                               64

TCGGACATGAAAGGACCCGTGTGAGCCGATTGCTATCTCCAGCGGCC      240
 S D M K G P C E P I V L S P A A                               80

CTGTATCCAGCTCACTCATCAATGGGCCAGTCAGGCCAGGCACTGG      288
 L S S S S L I N G P V R P R H W                               96

GCTCCGGAGGACTCACCAGTGCCTGCTGCCATGTGGACCTGGTGC      336
 A P E D S P L P P A A M W T W C                               112

AAGTTGAGGACTTCTTGTGGTCTAGTCACGCATGCAGTGTGGGGAT      384
 K L R T S C W S S H A C S V G D                               128

GCCTTGGTTTTTACTGCTCTGAGAATTGTTGAGATCTTACTAATAA      432
 A L V F T A L R I V E I L Y                               142

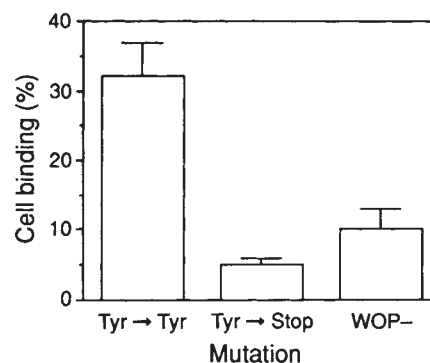
ACTGTGTAGTTGGAAAAA

```

FIG. 2 Sequence of *CAR*. Clone 27 was sequenced using the chain-termination method²⁵. The deduced open reading frame encodes 142 amino acids; an N-terminal myristoylation consensus sequence¹¹ is underlined, and the C-terminal tyrosine (shadowed) forms part of a putative tyrosine phosphorylation signature¹⁰ (bold).

FIG. 3 Inhibition of *CAR*-enhanced collagen adhesion by site-directed mutagenesis. Tyr 142 of *CAR* is essential for function: loss of this tyrosine by site-directed mutagenesis (Tyr → stop mutation) prevents the adhesion-inducing effect of *CAR* so that transfected WOP cells display only background levels of adhesion (as for untransfected WOP cell, WOP-), rather than the high levels in cells transfected with unchanged (Tyr → Tyr). $n=3$, mean $\pm$ s.e.m.

METHODS. Site-directed mutagenesis by polymerase chain reaction for 10 cycles was used. Primers: 5'-AATAGTACATGGGGAAGGTTATCA-3' with 5'-AATAGTACITATTAAAGTATCTCAACAA-3' for stop mutant, or with 5'-AATAGTACTTAGTAAAGTATCTCAAC-3' for tyrosine control. After ligation into pCDM8 and sequencing of the mutant and control plasmids to check the outcome of mutagenesis, the plasmids were independently electroporated into WOP cells which were assayed for collagen-binding ability, as in Fig. 2, after 48 h culture.



are on chromosome 5, $\alpha 3$ on chromosome 17 and $\beta 1$ is on chromosome 10 (ref. 15).

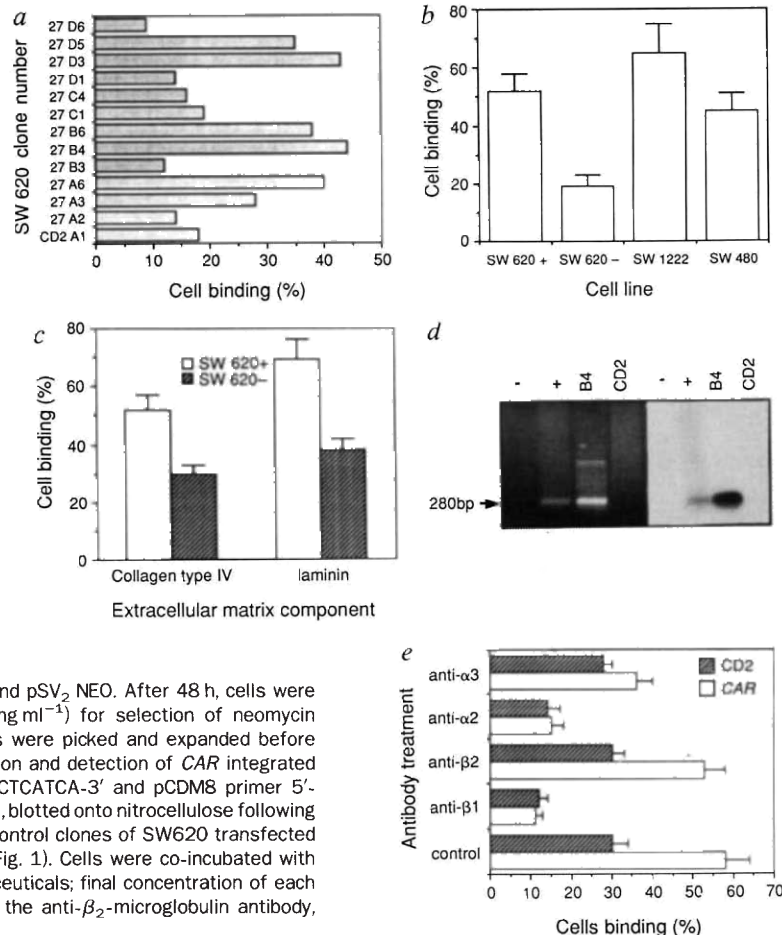
The effect of *CAR* on epithelial cell adhesion was studied by transfecting clone 27 into SW620, a colon cancer-derived line¹⁶. This caused two- to threefold enhanced binding to collagen type I, compared with the CD2-transfectants used as a control (background adhesion 15–20%; Fig. 4a). Cells of clone B4 of SW620, studied in detail (see below), showed 2.6-fold enhanced adhesion to collagen type I. This clone achieved adhesion levels greater than those of SW480, the primary tumour-derived line¹⁶ from which SW620 was derived through metastasis, and approached the levels of the cDNA library source, SW1222 cells (Fig. 4b). Binding of *CAR*-transfected SW620 to collagen type IV and laminin was also enhanced by similar amounts to collagen type I binding (Fig. 4c). Co-incubation of clone B4 *CAR*-transfected SW620 cells with antibodies against $\alpha 2$ and $\beta 1$ integrins reduced adhesion to collagen type I to the level

seen in untreated SW620 cells (Fig. 4e). A smaller reduction in adhesion was observed with anti- $\alpha 3$ integrin and no effect was seen with antibodies to lymphocyte $\beta 2$ integrin. No human integrin chains were detected in WOP cells transfected with clone 27 nor were there any major changes in the levels of expression of $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 5$, $\alpha 6$, $\beta 1$ or $\beta 4$ integrin chains in SW620 cells transfected with clone 27, by fluorescence-activated cell sorting analysis (data not shown).

CAR does not therefore seem to act by increasing expression of integrin heterodimers. Its structure and probable activation by tyrosine phosphorylation suggest a possible role in signal transduction, coupling the cytoplasmic portion of adhesion molecules to the cytoskeleton and enabling the external environment to influence cell shape and behaviour. These proposed mechanisms for controlling integrin binding, rather than levels of integrin expression, are consistent with several observations: carcinoma-derived cell lines express integrins $\alpha 2\beta 1$ and $\alpha 3\beta 1$

FIG. 4 *CAR* transfection of the epithelial line SW620 enhances cell adhesion to extracellular matrix. **a**, Independent clones of SW620 cells (numbered), transfected with *CAR*, show enhanced adhesion to collagen type I compared with control (CD2-transfected) cells ($n=2$). **b**, *CAR*-transfected SW620 cell clone B4 (+) showing two- to threefold enhanced adhesion to collagen type I compared with untreated (–) SW620, SW480 and SW1222 cells (mean $\pm$ s.e.m. for 6 experiments). **c**, Enhanced adhesion for collagen type IV and laminin in *CAR*-transfected cells (+; $n=3$). **d**, Stable integration of *CAR* into SW620 clone B4 cells, confirmed by Southern blotting of polymerase chain reaction product integration was also found only for those clones (A3, A6, B4, B6, D3, and D5; see **a**) that showed enhanced binding to extracellular matrix. Specific hybridization of a 280-bp product is shown for clone B4 and the positive control (+; clone 27 pCDM8) with absence of signal in the negative control (–; untreated SW620 cells) and CD2-transfected SW620 lanes. **e**, Augmented adhesion of *CAR*-transfected SW620 (clone B4, open bars) to collagen type I is inhibited by anti- $\alpha 2$ and anti- $\beta 1$ integrin antibodies. Antibodies to $\alpha 2$ or $\beta 1$ integrin reduced adhesion to the levels seen for CD2-transfected SW620 (hatched bars) antibodies to the $\alpha 3$ chain inhibited adhesion less. No inhibition was seen with antibodies to the lymphocyte $\beta 2$ chain ($n=3$, mean $\pm$ s.e.m.).

METHODS. SW620 cells (1×10^7) were co-transfected by electroporation (300 V) with clone 27-pCDM8 plasmid (15 μ g) and pSV2 NEO (1.5 μ g), both linearized with *Bam*HI. Controls were transfected with linearized pCDM8 containing CD2 (15 μ g) and pSV2 NEO. After 48 h, cells were cultured in DMEM with 10% FCS, supplemented with G418 (2 mg ml⁻¹) for selection of neomycin resistant clones. After 3–4 weeks of growth, independent clones were picked and expanded before screening for extracellular matrix adhesion. Primers for amplification and detection of *CAR* integrated into SW620 were: sequence-specific primer 5'-GTCATCCAGCTCACTCATCA-3' and pCDM8 primer 5'-CGCAGAACTGGTAGGTA-3'. Products were run on a 1.5% agarose gel, blotted onto nitrocellulose following established methods²³ and probed with ³²P-labelled clone 27. In control clones of SW620 transfected with CD2, expression was confirmed by surface fluorescence in (Fig. 1). Cells were co-incubated with anti- $\beta 1$, $\beta 2$, $\alpha 2$ and $\alpha 3$ monoclonal antibodies (Telios Pharmaceuticals; final concentration of each 1:1,000) in the 2-h adhesion assay. Control treatment was with the anti- $\beta 2$ -microglobulin antibody, BBM1 (ref. 26).



but do not bind to collagen type I (ref. 7); LFA-1 ligand binding avidity is modulated through the β -integrin subunit¹⁷; modulation of platelet integrin GPIIb-IIIa activity¹⁸ and the decrease in integrin function during keratinocyte differentiation¹⁹.

Loss of CAR activity could be a key early step in tumour invasion and metastasis, causing loss of differentiation induction or tumour cell release from cell or extracellular matrix attachments. The findings that SW620, the metastasis-derived line, displays lower binding to extracellular matrix than does SW480, its non-metastasis-derived partner from the same patient, and furthermore, that adhesiveness of SW620 to extracellular matrix is enhanced when transfected with CAR, support this notion. In addition, high levels (45–55%) of allelic loss on 16q, in the region of CAR, have been demonstrated in breast and prostatic cancers, possibly associated with later stage metastatic tumours^{20–22}. CAR is therefore a candidate tumour suppressor gene. □

Received 18 November 1991; accepted 24 January 1992.

1. Ruoslahti, E. & Giancotti, F. G. *Cancer Cells* **1**, 119–126 (1989).
2. Pignatelli, M. & Bodmer, W. F. *J. Path.* **162**, 95–97 (1990).
3. Bissel, M. J., Hall, G. H. & Parry, G. J. *Theor. Biol.* **99**, 31–68 (1982).
4. Takeichi, M. *Science* **251**, 1451–1455 (1991).
5. Ruoslahti, E. *J. Biol. Chem.* **264**, 13369–13372 (1989).
6. Hynes, R. O. *Cell* **48**, 549–554 (1987).
7. Pignatelli, M. P. & Bodmer, W. F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5561–5565 (1988).
8. Seed B. *Nature* **329**, 840–842 (1987).
9. Sternberg, M. J. E. *Comp. Applic. Biosci.* **7**, 257–260 (1991).
10. Patschinsky, T., Hunter, T., Esch, F. S., Cooper, J. A. & Sefton, J. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 973–977 (1982).
11. Towler, D. A., Gordon, J. I., Adams, S. P. & Glaser, L. A. *Rev. Biochem.* **57**, 69–99 (1988).
12. Resh, M. D. *Oncogene* **5**, 1437–1444 (1990).
13. Tapley, P., Horwitz, A., Buck, C., Duggan, K. & Rohrschneider, L. *Oncogene* **4**, 325–333 (1989).
14. Bicknell, D. C., Markie, D., Spurr, N. K. & Bodmer, W. F. *Genomics* **10**, 186–192 (1991).
15. Hemler, M. E. *A. Rev. Immun.* **8**, 365–400 (1990).
16. Leibovitz, A. *et al. Cancer Res.* **36**, 4562–4569 (1976).
17. Hibbs, M. L., Xu, H., Stacker, S. A. & Springer, T. A. *Science* **251**, 1611–1613 (1991).
18. Du, X. *et al. Cell* **65**, 409–416 (1991).
19. Adams, J. C. & Watt, F. *Cell* **63**, 425–435 (1990).
20. Larsson, C., Bystrom, C., Skoog, L., Rotstein, S. & Nordenskjold, M. *Genes Chrom. Cancer* **2**, 191–197 (1990).
21. Sato, T. *et al. Cancer Res.* **50**, 7184–7189 (1990).
22. Bergerheim, U., Kunimi, K., Collins, V. & Ekman, P. *Genes Chrom. Cancer* **3**, 215–220 (1991).
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning, A Laboratory Manual* 2nd edn (Cold Spring Harbor Laboratory Press, New York, 1989).
24. Huscchtscha, L. I., Lucibello, F. C. & Bodmer, W. F. *In Vitro Cell. dev. Biol.* **25**, 105–108 (1989).
25. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
26. Brodsky, F. M., Bodmer, W. F. & Parham, P. *Eur. J. Immun.* **9**, 536–545 (1979).

ACKNOWLEDGEMENTS. We thank H. Durbin for help in the early phases of this work, and P. Bates for his help with interpreting the CAR sequence.

SEC65 gene product is a subunit of the yeast signal recognition particle required for its integrity

Byron C. Hann, Colin J. Stirling* & Peter Walter

Department of Biochemistry and Biophysics, University of California, Medical School, San Francisco, California 94143-0448, USA

* Department of Biochemistry and Molecular Biology, University of Manchester, School of Biological Sciences, Manchester M13 9PT, UK

PROTEIN targeting to the endoplasmic reticulum (ER) in mammalian cells is catalysed by the signal recognition particle (SRP), which consists of six protein subunits and an RNA subunit^{1,2}. *Saccharomyces cerevisiae* SRP is a 16S particle, of which only two subunits have been identified: a protein subunit, SRP54p, which is homologous to the mammalian SRP54 subunit, and an RNA subunit, scr1 (ref. 3). The *sec65-1* mutant yeast cells⁴ are temperature-sensitive for growth and defective in the translocation of several secreted and membrane-bound proteins. The DNA sequence of the SEC65 gene suggests that its product is related to mammalian SRP19 subunit and may have a similar function⁵.

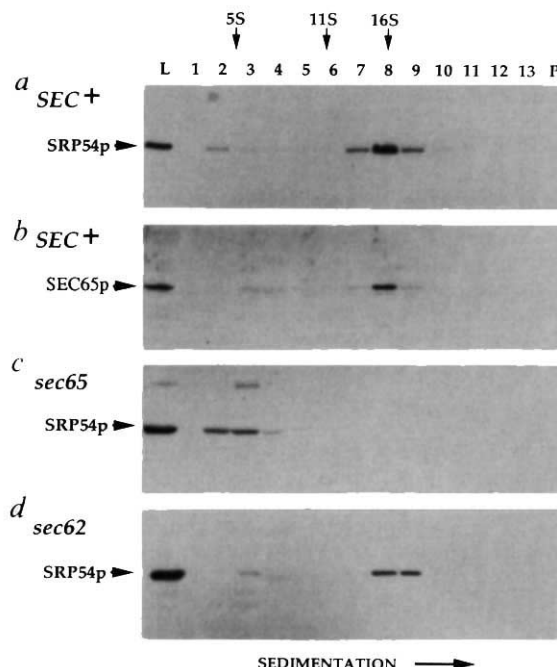


FIG. 1 SEC65p cosediments with *S. cerevisiae* SRP. Crude cell extracts prepared from wild-type (a and b), *sec65-1* (c) or *sec62-1* cells (d) were layered over a 5–20% sucrose gradient and centrifuged for 15 h at 40,000 r.p.m. Fractions from the gradient (1–13), load (L) and pellet (P) were analysed, after SDS-PAGE, on western blots developed with either affinity-purified anti-SRP54p (ref. 3; a, c, d) or anti-SEC65p (ref. 5; b). Western blot analysis of sucrose gradient fractions from a wild-type strain isogenic to *sec65-1* developed with anti-SRP54p were indistinguishable from the blot shown in b (data not shown). Sedimentation values of protein standards are indicated.

METHODS. The yeast cell extracts and sucrose gradients were prepared as previously described³. The *sec65-1* and *sec62-1* cell extracts were prepared from cells grown at 24 °C. Bound antibodies were visualized on the blots by ¹²⁵I-labelled second antibody (a, c and d) or ECL (Amersham; b). Anti-SEC65p was raised against a fusion protein⁵.

Here we show that SEC65p is a subunit of the *S. cerevisiae* SRP and that it is required for the stable association of another subunit, SRP54p, with SRP. Overexpression of SRP54p suppresses both growth and protein translocation defects in *sec65-1* mutant cells.

To determine whether SEC65p is a component of the *S. cerevisiae* SRP, we first fractionated total cell extracts prepared from wild-type cells by sucrose-density centrifugation. Gradient fractions were analysed on western blots developed with anti-

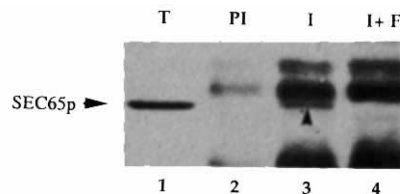
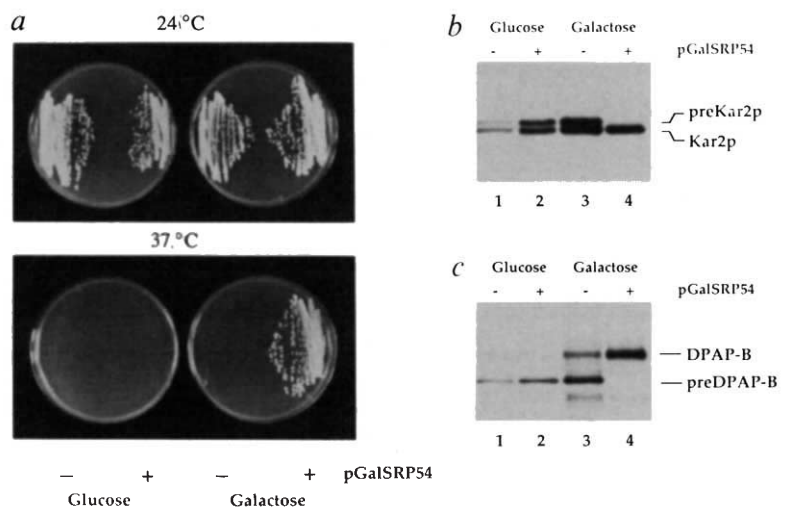


FIG. 2 SEC65p coimmunoprecipitates with SRP54p. Total yeast cell extracts (T, lane 1) and the products of native immunoprecipitations from these extracts, performed with either preimmune serum (PI, lane 2) or anti-SRP54p immune sera in the absence (I, lane 3) or presence of SRP54p-glutathione transferase fusion protein (I+F, lane 4), were examined on western blots developed with anti-SEC65p. Arrowhead indicates SEC65p (lane 3).

METHODS. Native immunoprecipitations were done from 100 μ l of crude cell extract (protein concentration ~ 30 mg ml⁻¹) using Protein-A-Sepharose complexed with anti-SRP54p³. The western blot spans a relative molecular mass range of 40,000 to 27,000. Background bands are derived from the heavy and light chains of the first antibody.

FIG. 3 Overexpression of SRP54p suppresses growth and translocation defects of *sec65-1* cells. **a**, Untransformed *sec65-1* cells or cells transformed with pGalSRP54 (ref. 3) were streaked on either YEP-Dex or YEP-Gal plates and grown at either 24 °C or 37 °C. The growth rate of cells containing the plasmid on YEP-Gal was comparable to that of the isogenic wild-type strain (data not shown). **b**, **c**, After growth for 1 h at the nonpermissive temperature (37 °C), untransformed *sec65-1* cells (lanes 1, 3) or cells transformed with pGalSRP54 (lanes 2, 4) were pulse-labelled in either glucose-containing (lanes 1, 2) or galactose-containing media (lanes 3, 4). Cell extracts were prepared and non-native immunoprecipitations were done using anti-Kar2p (**b**) or anti-DPAP-B (**c**) antibodies. Mature (Kar2p, DPAP-B) and precursor (preKar2p, preDPAP-B) forms of Kar2p and dipeptidyl aminopeptidase-B are indicated.

METHODS. **a**, The *sec65-1* cells were transformed using the lithium acetate method⁸. **b**, **c**, The *sec65-1* cell were grown to log phase in YEP-Dex or YEP-Gal liquid medium at 24 °C, shifted to 37 °C for 45 min and pelleted. Cells were resuspended in prewarmed minimal medium plus supplements, and grown for an additional 15 min at 37 °C. Pulse-labelling was done with ³⁵S-methionine (ICN) for 10 min. Cell extraction and non-native immunoprecipitation was performed as described³.



bodies raised against SRP54p (Fig. 1a) and SEC65p (Fig. 1b). SEC65p cosediments with SRP54p, which is consistent with the notion that both proteins are components of the same particle. To demonstrate directly that both proteins are SRP subunits, we immunoprecipitated SRP under native conditions with anti-SRP54p. We analysed the precipitated fraction on western blots developed with anti-SEC65p antibodies. SEC65p is immunoprecipitated with anti-SRP45p serum (Fig. 2, lane 3), but not with preimmune serum (lane 2). An excess of SRP54p (added as a fusion protein) prevented precipitation of SEC65p (Fig. 2, lane 4).

Mammalian SRP19 is required for the association of SRP54 with SRP RNA⁶⁻⁸. We therefore asked whether SRP is physically altered in *sec65-1* cells. Cell extracts prepared from *sec65-1* cells were fractionated on sucrose-density gradients and the fractions analysed on western blots developed with anti-SRP54p antibodies. The bulk of SRP54p was seen to sediment as an apparently monomeric protein near the top of the gradient (Fig. 1c). This effect was not a consequence of presecretory protein accumulation in the mutant cells as SRP54p remained part of a 16S particle in cell extracts prepared from another secretory mutant strain, *sec62-1* (Fig. 1d). In western blots of extracts of *sec65-1* cells, the concentration of SEC65p was diminished and after sucrose gradient fractionation the protein could no longer be detected (data not shown). This is consistent with the finding that the concentration of the mutant protein is reduced after a shift to 37 °C (ref. 5). We conclude that in *sec65-1* cells the interaction of SRP54p with the rest of the particle has been disrupted. SEC65p is therefore required for the integrity of SRP.

We next looked for genetic interactions between these two SRP subunits. Overexpression of SRP54p suppressed the temperature-sensitive growth phenotype of *sec65-1* cells (Fig. 3a). The *sec65-1* cells containing a plasmid-borne copy of the *SRP54* gene under the control of the *GAL1* promoter (pGalSRP54) were able to grow at 37 °C, the nonpermissive temperature, on galactose but not on glucose-containing media. These results are supported by the independent isolation of a multicopy

suppressor of the *sec65-1* defect, which has been identified as *SRP54* (ref. 4).

If the overproduction of SRP54p acts directly to rescue *sec65-1* cell growth at 37 °C, then the protein translocation defect in these cells should also be suppressed. To test if this is so, we immunoprecipitated proteins that use of the endoplasmic reticulum translocation system from cell extracts of *sec65-1* cells. Precursor forms of both a soluble protein, Kar2p, and a membrane protein, DPAP-B, were detected in *sec65-1* cells which did not harbour pGalSRP54 (Fig. 3b, c; lanes 1, 3) as well as in *sec65-1* cells transformed with pGalSRP54 and grown on glucose (Fig. 3b, c; lane 2). In cells overexpressing SRP54p, however, the untranslocated precursor forms of the proteins were no longer detected (Fig. 3b, c; lane 4). The overexpression of SRP54p, therefore, efficiently suppresses both the growth defect and the protein translocation defect of *sec65-1* cells.

These results agree with biochemical studies of mammalian SRP, which suggest that SRP19 and SRP54 interact during SRP assembly. Because the binding of mammalian SRP54 to SRP RNA requires the presence of SRP19, it has been proposed that SRP19 functions to stabilize a conformation of SRP RNA that forms a high affinity binding site for SRP54 (refs 7, 8). Similarly, yeast SEC65p could function to stabilize a conformation of scR1 that contains a high-affinity binding site for SRP54p. We have shown that the affinity of SRP54p for the rest of SRP is reduced in *sec65-1* cells. High concentrations of SRP54p could drive the formation of a functional SRP by promoting binding of SRP54p to a reduced-affinity binding site and consequently suppress the *sec65-1* mutant phenotype. Two observations⁵, however, indicate that the interaction between SEC65p and SRP54p may be more complex. First, the overproduction of SRP54p can stabilize mutant SEC65p in *sec65-1* cells, and second, such overproduction appears insufficient to bypass a requirement for SEC65p in cells in which SEC65 has been deleted. Thus the assembly of SRP54p onto scR1 appears to require SEC65p, and SEC65p and SRP54p could interact directly and/or cooperatively in SRP. □

Received 15 November 1991; accepted 29 January 1992.

- Walter, P. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7112-7116 (1980).
- Walter, P. & Blobel, G. *Nature* **99**, 691-698 (1982).
- Hann, B. C. & Walter, P. *Cell* **67**, 131-143 (1991).
- Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R. & Schekman, R. *Molec. Biol. Cell* (in the press).
- Stirling, C. J. & Hewitt, E. W. *Nature* **356**, 534-537 (1992).
- Walter, P. & Blobel, G. *Cell* **34**, 525-533 (1983).
- Zopf, D., Bernstein, H. D., Johnson, A. E. & Walter, P. *EMBO J.* **9**, 4511-4517 (1990).

- Rörmisch, K., Webb, J., Lingelbach, K., Gausepohl, H. & Dobberstein, B. *J. cell. Biol.* **111**, 1793-1802 (1990).
- Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bact.* **153**, 163-168 (1983).

ACKNOWLEDGEMENTS. We thank T. Stevens (Univ. of Oregon) and M. Rose (Princeton University) for providing anti-DPAP-B and anti-Kar2p antisera; R. Deshaies and the members of the Walter laboratory for comments on the manuscript. This work was supported by grants from the NIH and the Alfred P. Sloan Foundation.

The *S. cerevisiae* SEC65 gene encodes a component of yeast signal recognition particle with homology to human SRP19

Colin J. Stirling & Eric W. Hewitt

Department of Biochemistry and Molecular Biology, Medical School, University of Manchester, Manchester M13 9PT, UK

TRANSLOCATION of proteins across the endoplasmic reticulum (ER) membrane represents the first step in the eukaryotic secretory pathway. In mammalian cells, the targeting of secretory and membrane protein precursors to the ER is mediated by signal recognition particle (SRP), a cytosolic ribonucleoprotein complex comprising a molecule of 7SL RNA and six polypeptide subunits (relative molecular masses 9, 14, 19, 54, 68 and 72K)¹. In *Saccharomyces cerevisiae*, a homologue of the 54K subunit (SRP54)^{2,3} co-purifies with a small cytoplasmic RNA, scR1 (refs 4, 5). Genetic data indicate that SRP54 and scR1 are involved in translocation *in vivo*, suggesting the existence of an SRP-like activity in yeast^{5,6}. Whether this activity requires additional components similar to those found in mammalian SRP is not known. We have recently reported a genetic selection that led to the isolation of a yeast mutant, *sec65-1*, which is conditionally defective in the insertion of integral membrane proteins into the ER⁷. Here we report the cloning and sequencing of the *SEC65* gene, which encodes a 31.2K protein with significant sequence similarity to the 19K subunit of human SRP (SRP19)⁸. We also report the cloning of a multicopy suppressor of *sec65-1*, and its identification as the previously defined *SRP54* gene, providing genetic evidence for an interaction between these gene products *in vivo*.

The wild-type *SEC65* gene was cloned from a YEp24-based multicopy plasmid library, by complementation of the temperature-sensitive growth phenotype of the *sec65-1* mutant strain CSY126. The plasmid p65-B1 contains a 13.5-kb genomic fragment of which 2.2 kb fully complements the *sec65-1* defect when cloned in either single- or multicopy plasmid vectors (Fig. 1). Integrative mapping has shown that p65-B1 contains the authentic *SEC65* gene. In summary, the genomic sequences in p65-B1 have been used to direct integration of a selectable marker (*URA3*) into the yeast genome by homologous recombination⁹. In subsequent genetic crosses the integrated *URA3* marker exhibits 100% linkage to the *sec65-1* locus, demonstrating that the site of insertion is at the genomic *SEC65* locus, and that p65-B1 represents a clone of that locus.

The nucleotide sequence of the 2,219 base pairs (bp) minimum complementing fragment from p65-B1 contains a long open reading frame (340–1,159bp), which would encode a 273-amino-acid polypeptide with a predicted *M_r* of 31,170 (Fig. 2). The complementation data (Fig. 1) indicates that this open reading frame represents the *SEC65* gene. A second truncated reading frame was noted which terminates at 1,577bp and extends beyond 2,219bp. This represents codons 13–226 of the *URA5* gene (see Fig. 2 legend). A comparison of the predicted amino-acid sequence of the SEC65 protein (SEC65p) with sequence databases identified a significant match to human SRP19. The optimized alignment reveals 28% sequence identity along the entire length of SRP19 (144 amino acids), with an overall 50% similarity when conservative substitutions are included. The alignment of SRP19 with the C-terminal portion of SEC65p suggests the existence of an additional N-terminal domain in the yeast protein.

Despite the fact that the *sec65-1* mutation confers a conditional-lethal phenotype, we have found *SEC65* to be a non-essential gene. A targeted gene disruption has been constructed in which 70% of the *SEC65* coding sequence has been replaced

by a functional copy of the *HIS3* gene (Fig. 3a, b). The resulting *sec65::HIS3* null mutant (CSY186) is temperature sensitive for growth at 37 °C, indicating that *SEC65* is essential for growth at the elevated temperature. In other respects the *sec65* null mutant exhibits a phenotype similar to targeted gene disruptions in either *SRP54* or *SCR1* (refs 4, 5), an observation consistent with all three genes encoding components of the same functional complex. All three gene disruptions result in very slow growing strains, which exhibit a 'petite' phenotype, suggesting a defect in mitochondrial function; whether this indicates a direct role for yeast SRP in mitochondrial protein import is unknown. As expected CSY186 cells exhibit a defect in the translocation of both secretory and membrane protein precursors into the ER (Fig. 3c, d). The defect is markedly less severe than that seen in *sec65-1* cells which, at 37 °C, accumulate >90% of dipeptidyl-aminopeptidase-B (DPAPB) in a precursor form⁷. CSY186 cells are temperature sensitive for growth, yet there is no significant change in the severity of translocation defect observed in these cells at 37 °C (not shown).

This anomaly between *sec65-1* and *sec65::HIS3* could be due to a deleterious effect exerted by the mutant protein at 37 °C, but the recessive nature of the *sec65-1* mutation makes this seem unlikely. Moreover, the temperature-sensitive defect in CSY126 cells correlates with a near total loss of detectable SEC65p antigen in western blots (Fig. 4). An apparently related anomaly has been noted for SRP54p; the immediate consequences of its *in vivo* depletion include a severe translocation defect which is diminished after prolonged propagation of the *srp54* null mutant⁵. These data suggest that the null mutants have adapted to the loss of SRP function, perhaps via the induction of a bypass/salvage pathway; cytosolic HSP70s could be involved in this phenomenon⁵. The slow growing phenotype of SRP-null mutants, coupled with this apparently adaptive response, suggests that although SRP is not essential for translocation in yeast, it is important in the normal ER-targeting pathway in this organism.

In addition to the authentic *SEC65* gene, we have isolated an extragenic multicopy suppressor of the *sec65-1* mutation, p65-A6. The genomic DNA in p65-A6 does not suppress *sec65-1* when present on a single-copy plasmid vector, and it directs the integration of a linked *URA3* marker to a site in the yeast genome which is unlinked to *sec65-1*. In a separate study¹⁰, we

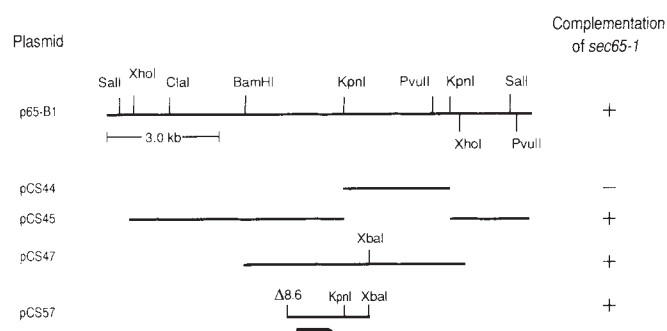


FIG. 1 Defining the minimum complementing sequence within p65-B1. Top line represents a restriction enzyme map of the genomic DNA inserted in p65-B1, which includes the *SEC65* open reading frame (solid arrow). A series of subclones (left) have been constructed and tested for their ability to complement the temperature-sensitive growth defect of strain CSY126 (*sec65-1*, *ura3-52*, *leu2-3-112*, *MATa*), at 37 °C (right).

METHODS. Plasmid pCS44 corresponds to the 4-kb *KpnI* fragment from p65-B1 inserted in to the multicopy vector YEp352 (ref. 11). Plasmid pCS45 was generated by deleting the 4-kb *KpnI* fragment from p65-B1. pCS47 represents the 6.8-kb *BamHI/XhoI* fragment from p65-B1 inserted into the *BamHI/SalI* sites of YEp352. DNA inserted into pCS57 corresponds to an *ExoII* deletion derivative of pCS51 (pCS51Δ8.6; described in Fig. 2). The insert from pCS51Δ8.6 was transferred as an *XbaI/PvuII* fragment into the *XbaI/SmaI* site of the single-copy vector pRS316 (ref. 12).

have shown that the *sec65-1* mutant defect is completely suppressed by overexpression of the yeast *SRP54* gene. Indeed, the restriction map of the complementing region from p65-A6 is very similar to that predicted from the nucleotide sequence of the yeast *SRP54* gene (not shown)². Furthermore, in southern blots, *SRP54* sequences hybridize at high stringency to the expected portion of the p65-A6 DNA, and to a single gene copy in yeast genomic DNA² (not shown). We therefore conclude that the multicopy suppressor of *sec65-1* is *SRP54*.

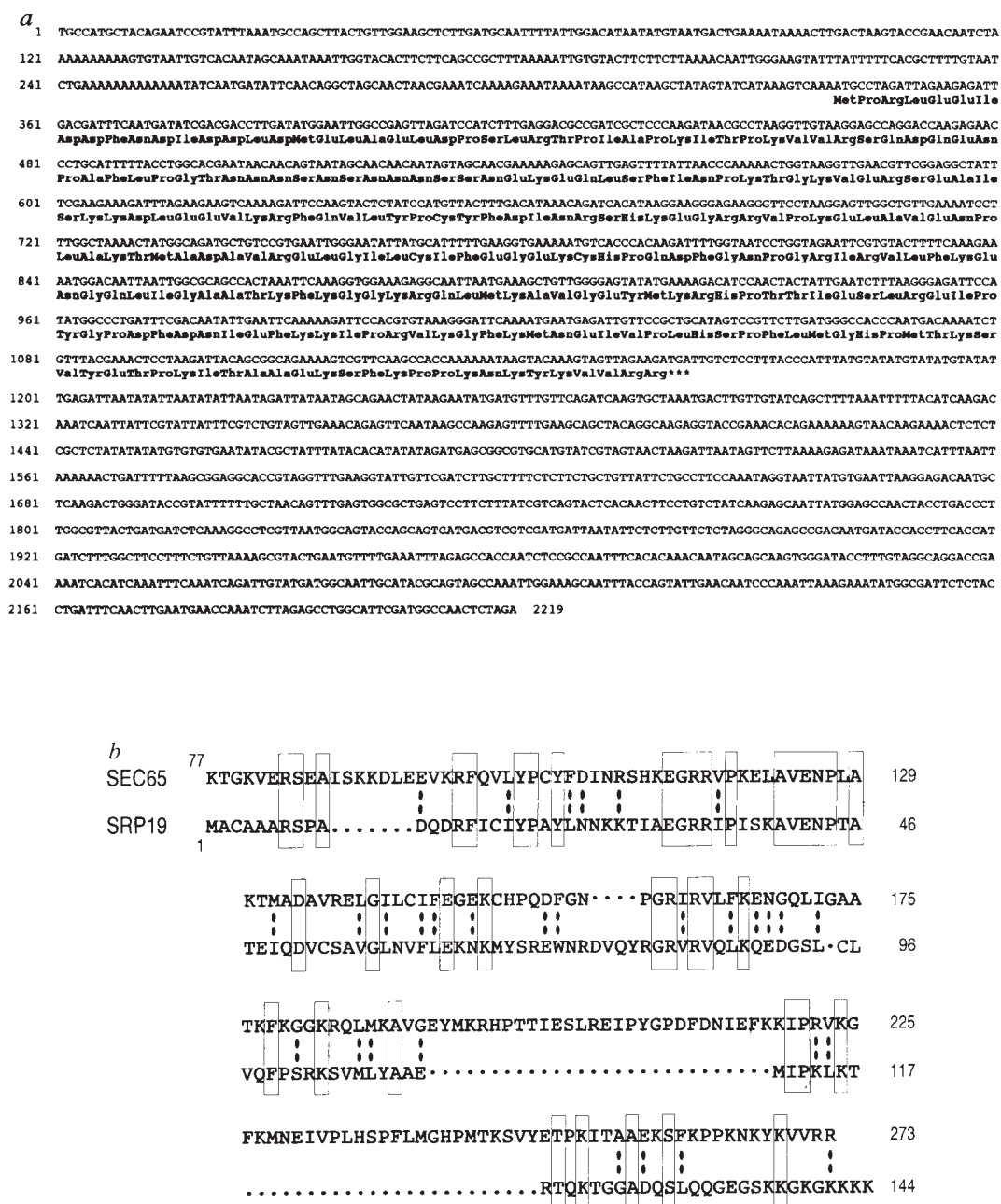
Multicopy *SRP54* (p65-A6) does not suppress the growth defect in the *sec65::HIS3* null mutant (not shown). Therefore, the observed suppression of *sec65-1* cannot be due to a simple bypass of the SEC65p-dependent step, but requires an interaction between SRP54p and the mutant sec65 protein (sec65p). A direct interaction is consistent with the observation that p65-A6 partially stabilizes the level of sec65p observed in CSY126 cells

at 37 °C (Fig. 4). In the light of these data we propose the following model for the mechanism of multicopy suppression of *sec65-1*. The reduction in sec65p concentration at 37 °C could be due to the rapid turnover of a thermolabile mutant protein. This would then limit the ability of sec65p to interact with SRP54p, perhaps exacerbated by a reduced affinity for SRP54p. Under these circumstances, the overexpression of SRP54p could promote the formation of an SRP54p/sec65p-containing complex in which the sec65p moiety is effectively stabilized.

The data presented here identify the *SEC65* gene product as the yeast homologue of human SRP19, and represents the first identification of a component of SRP via the isolation of a translocation-defective mutant. The suppression of *sec65-1* by *SRP54* provides genetic evidence for an interaction between these gene products *in vivo*. In the accompanying paper¹⁰, we demonstrate that yeast SRP is comprised of at least three

FIG. 2 Nucleotide and deduced amino-acid sequence of *SEC65*. *a*, Nucleotide sequence of the 2,219-bp Δ 8.6-*XbaI* fragment is shown, with the deduced amino-acid sequence shown below the coding sequence. The *SEC65* reading frame extends from the ATG start codon at 340 bp, to a TGA stop codon at 1,159 bp. The truncated *URA5* reading frame, 2,219–1,577 bp, is not shown. Sequences from 968–2,219 bp overlap with, and differ from, those published by de Montigny *et al.*¹³ Of two base substitutions in the *URA5* coding sequence, one is silent, and the other results in a conservative amino-acid change from Ser to Asp at residue 138. There are nine insertions/deletions outside the *URA5* coding sequence. Six of these disrupt the *SEC65* reading frame, and we believe the previously published sequence is incorrect in this region. *b*, Alignment of SEC65p and SRP19 amino-acid sequences. Identical residues are boxed, with chemically similar residues denoted by double dots. Gaps introduced to maximize alignment are indicated by single dots.

METHODS. The 3.5-kb *Bam*HI/*Xba*I fragment containing the *SEC65* gene was inserted into phagemid vectors pUC118 and pUC119 (ref. 14), generating pCS52 and pCS51 respectively. Complementary sets of nested *Exo*III deletions were generated¹⁵, using *Bam*HI/*Sst*I cut pCS51 and *Sph*I/*Xba*I cut pCS52 as substrates for *Exo*III digestion. Single-stranded sequencing templates¹⁴ were sequenced using the Sequenase II modification (USB Corp.) of the dideoxy-chain termination method.



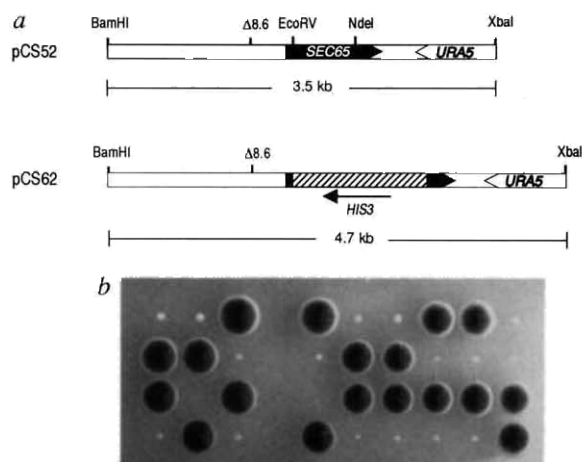
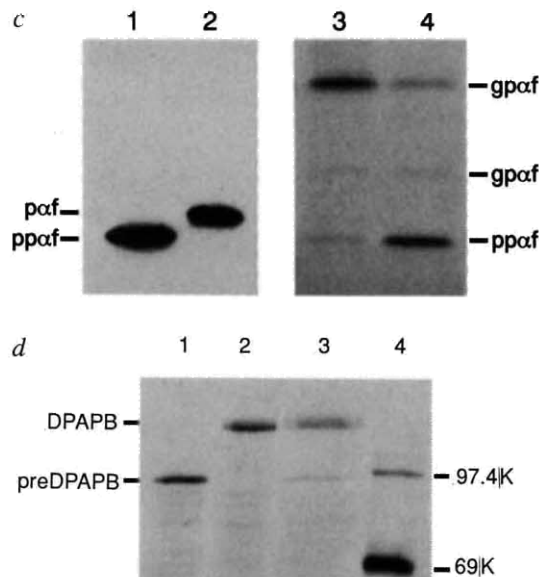


FIG. 3 Disruption of the *SEC65* gene results in slow-growing cells that exhibit defects in the translocation of ER proteins. *a*, Construction of a disrupted *sec65* allele. 3.5-kb *Bam*HI/*Xba*I fragment, cloned in pCS52 (top) is shown with the *SEC65*, and *URA5* reading frames indicated by arrows. The *sec65::HIS3* allele was constructed by inserting *HIS3* (ref. 16) as a 1.8-kb *Bam*HI (blunt) fragment into the *Eco*RV/*Nde*I sites of pCS52, to produce pCS62 (bottom). *HIS3* sequences are depicted by a hatched box, with the *HIS3* reading frame indicated by an arrow¹⁶. *b*, The 4.7-kb *Bam*HI/*Xba*I fragment from pCS62 was used to transform the diploid strain TR1 (genotype listed below)¹⁷ to *His*⁺. Stable *His*⁺ transformants were sporulated, asci dissected, and the phenotypes of haploid progeny examined. The result for one such strain, CSY181 is shown. Nine tetrads were dissected, with the four haploid spores from each tetrad arranged vertically. Generally, all four spores are viable, but segregate 2:2 as large and small colonies. Small colonies carry the *sec65::HIS3* allele, large colonies are *His*⁺ and carry a non-disrupted *SEC65* gene. The structure of the integrated locus in CSY181, and its segregation amongst haploids, has been confirmed by Southern blotting (not shown). The minimum complementing plasmid, pCS57, fully suppresses the null-mutant phenotype (after sporulation of a plasmid-containing diploid), demonstrating that the observed phenotype is specifically due to the loss of *SEC65* (not shown). CSY186, carrying the *sec65::HIS3*



gene disruption, accumulates translocation precursors of both pre-pro- α -factor (*c*), and DPAPB (*d*). *c*, *in vitro* translated pre-pro- α -factor (lane 1), immunoprecipitated forms of α -factor from CSY186 cells (lane 4), or from wild-type cells labelled in the presence (lane 2), or absence (lane 3) of tunicamycin ($10 \mu\text{g ml}^{-1}$). Pre-pro- α -factor (pp α f), pro- α -factor (p α f), and core glycosylated pro- α -factor (gp α f) are indicated (pp α f migrates more rapidly than p α f in this gel system). *d*, Immunoprecipitated forms of DPAPB from CSY186 (lane 3), or from wild-type cells labelled in the presence (lane 1), or absence (lane 2) of tunicamycin, molecular weight standards (lane 4). Mature and unglycosylated precursor forms are labelled DPAPB and preDPAPB respectively.

METHODS. Strains CSY186 and TR3 (below) were pulse-labelled (5 min at 30°C) with Tran^{35}S -label, whole cell extracts prepared, then specific polypeptides immunoprecipitated and resolved by SDS-PAGE as described previously⁷. Genotypes: TR1 (*trp1/trp1, lys2/lys2, his3/his3, ura3/ura3, ade2/ade2, MATa/MATa*), CSY181 (as TR1 but *SEC65/sec65::HIS3*). TR3 (*trp1, lys2, his3, ura3, ade2, MATa*), CSY186 (as TR3 but *sec65::HIS3*).

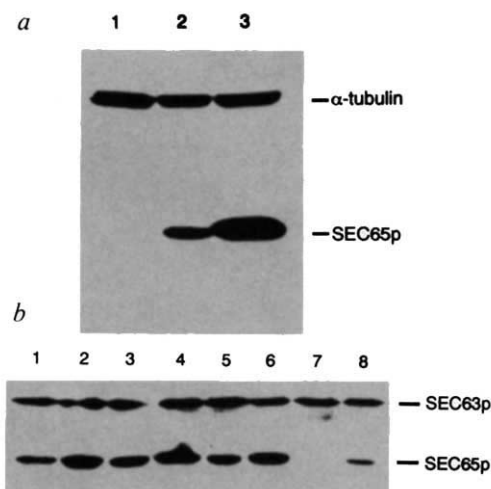


FIG. 4 Identification of the *SEC65* protein. Antibodies raised against a protein-A/*SEC65*p fusion-protein recognize a single polypeptide species in immunoblots of whole-cell yeast extracts. The M_r of this species, 32K, corresponds closely to that predicted from the deduced amino-acid sequence (Fig. 2). This 32K protein is not recognized by pre-immune serum (not shown). *a*, Western blot of whole cell extracts prepared from CSY186 (*sec65::HIS3*; lane 1), TR1 + YEp24 (wildtype; lane 2), or TR1 + p65-B1 (multicopy *SEC65*; lane 3), probed with both anti-*SEC65*p antiserum, and TAT1 monoclonal antibody (specific for α -tubulin)¹⁸, as indicated; α -tubulin serves as the loading control. Only relevant regions of the blot are shown. *b*, The level of *sec65* protein in CSY126 (*sec65-1*) is dramatically reduced after a shift to 37°C . Whole-cell extracts were prepared from cultures grown continually at 24°C (lane 1–4), or after a 4 h shift to 37°C (lanes 5–8): TR3 (lanes 1 and 5), CSY126 + pCS57 (*SEC65*⁺; 4b, lanes 2 and 6), CSY126 + pRS316 (lanes 3 and 7), or CSY126 + p65-A6 (lanes 4 and 8). Parallel Western blots were probed with anti-*SEC65*p, or anti-*SEC63*p antiserum (D. Feldheim, unpublished) as indicated. *SEC63*p signal serves as a control.

METHODS. Strains were grown in minimal medium, selecting for the appropriate plasmid markers. Whole cell extracts were prepared by glass bead lysis, resolved by SDS-PAGE (12.5% polyacrylamide gel), and transferred to nitrocellulose⁷. Each lane was loaded with extract from an equivalent of 0.5 OD₆₀₀ units of cells. Bound antibodies were detected using the ECL system (Amersham). The anti-*SEC65*p antiserum was raised against a protein-A/*SEC65*p fusion protein as follows. The protein-A-*SEC65* gene-fusion was constructed by inserting the 1.8-kb *Eco*RV/*Xba*I fragment from pCS52 into the protein-A fusion vector pRIT33 (*Sma*I/*Xba*I; a derivative of pRIT 2T)¹⁹, to generate pEH1. High-level expression of pEH1 fusion protein was induced with isopropyl- β -D-thiogalactopyranoside. The fusion protein was purified to homogeneity by affinity chromatography of whole-cell extracts on IgG-sepharose²⁰. Purified material was used to immunize rabbits. Genotypes: CSY126 (*sec65-1, ura3-52, leu2-3,-112, MATa*).

subunits, SEC65p, SRP54p and scR1, and furthermore that SEC65p, like its mammalian counterpart, is required for the stable association of SRP54p with this complex. □

Received 15 November 1991; accepted 29 January 1992.

- Walter, P. & Blobel, G. *Nature* **99**, 691–698 (1982).
- Hann, B. C., Poritz, M. A. & Walter, P. *J. Cell Biol.* **109**, 3223–3230 (1989).
- Amaya, Y., Nakano, A., Ito, K. & Mori, M. *J. Biochem.* **107**, 457–463 (1990).
- Felici, F., Cesareni, G. & Hughes, J. M. *Molec. Cell Biol.* **9**, 3260–3268 (1989).
- Hann, B. C. & Walter, P. *Cell* **67**, 131–144 (1991).
- Amaya, Y. & Nakano, A. *FEBS Lett.* **283**, 325–328 (1991).
- Stirling, C. J. *Molec. Biol. Cell* (in the press).
- Lingelbach, K. et al. *Nucleic Acids Res.* **16**, 9431–9442 (1988).
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354–6358 (1981).

- Hann, B. C., Stirling, C. J. & Walter, P. *Nature* **356**, 532–533 (1992).
- Hill, J. E., Myers, A. M., Koerner, T. J. & Tzagoloff, A. *Yeast* **2**, 163–168 (1986).
- Sikorski, R. S. & Hieter, P. *Genetics* **122**, 19–27 (1989).
- de Montigny, J., Belarbi, A., Hubert, J. C. & Lacroute, F. *Molec. gen. Genet.* **215**, 455–462 (1989).
- Vieira, J. & Messing, J. *Meth. Enzym.* **153**, 3–11 (1987).
- Henikoff, S. *Meth. Enzym.* **155**, 156–165 (1987).
- Struhl, K. *Nucleic Acids Res.* **13**, 8587–8601 (1985).
- Parker, R. et al. *Molec. Cell Biol.* **8**, 3150–3159 (1988).
- Woods, A. et al. *J. Cell Sci.* **93**, 491–500 (1989).
- Nilsson, B. L., Abrahmsen, L. & Uhlen, M. *EMBO J.* **4**, 1075–1080 (1985).
- Deshaies, R. J. & Schekman, R. *Molec. Cell Biol.* **10**, 6024–6035 (1990).

ACKNOWLEDGEMENTS. We thank D. Feldheim, K. Gull and T. Stevens for providing antibodies; J. Rine for providing a YEp24-based genomic library; E. Doherty for technical assistance; and N. Bulleid, D. Hardy, D. Thornton and P. Woodman for reading the manuscript. We also acknowledge our use of the SERC Sequencing facility for sequence manipulation and analysis. This work was supported by the SERC, the Royal Society, the Society for General Microbiology, and the Nuffield Foundation. E.W.H. holds a SERC studentship.

Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code

J. D. Bain*, Christopher Switzer†, A. Richard Chamberlin* & Steven A. Benner‡

* Department of Chemistry, University of California, Irvine, California 92717, USA

† Department of Chemistry, University of California, Riverside, California 9251, USA

‡ Laboratory for Organic Chemistry, Eidgenössische Technische Hochschule, CH-8092 Zurich, Switzerland

ONE serious limitation facing protein engineers is the availability of only 20 'proteinogenic' amino acids encoded by natural messenger RNA. The lack of structural diversity among these amino acids restricts the mechanistic and structural issues that can be addressed by site-directed mutagenesis. Here we describe a new technology for incorporating non-standard amino acids into polypeptides by ribosome-based translation. In this technology, the genetic code is expanded through the creation of a 65th codon-anticodon pair from unnatural nucleoside bases having non-standard hydrogen-bonding patterns^{1,2}. This new codon-anticodon pair efficiently supports translation *in vitro* to yield peptides containing a non-standard amino acid. The versatility of the ribosome as a synthetic tool offers new possibilities for protein engineering, and compares favourably with another recently described approach in which the genetic code is simply rearranged to recruit stop codons to play a coding role^{3–9}.

To compare the alternative strategies of rearrangement and expansion of the genetic code for increasing the range of amino acids that can be incorporated into proteins by translation, two mRNA molecules were prepared (Fig. 1). Both encode identical hexadecapeptides that are preceded by identical consensus 5'-

untranslated sequences¹⁰. One, however, has the UAG nonsense (stop) codon at position 9 as the signal for the incorporation of the non-standard amino acid L-iodotyrosine, whereas the other uses the novel non-standard (*iso*-C)AG codon (the 65th codon) at this position (Fig. 2) for the same purpose. These mRNA molecules were incubated separately with rabbit reticulocyte lysate containing L-[³H]leucine and L-[³⁵S]methionine. Ribosome-mediated peptide synthesis was evaluated in the presence and absence of the corresponding charged and uncharged transfer RNA molecules incorporating either CUA or CU(*iso*-dG) as the anticodon, where *iso*-G is a non-standard purine complementary to *iso*-C^{11,12} (Fig. 2). Translation products were isolated by precipitation (Table 1) or high-performance liquid chromatography (Table 2) and identified by comparison with authentic standards¹³.

A 'suppression level' of 63% (Table 1) or 67% (Table 2) was observed for read-through of a UAG nonsense codon in the presence of a semi-synthetic suppressor tRNA incorporating the CUA anticodon and charged with iodotyrosine. In contrast, read-through of the (*iso*-C)AG codon was 90% (Table 1) or 91% (Table 2) in the presence of the corresponding non-standard tRNA_{CU(*iso*-dG)} charged with iodotyrosine.

The specificity of translation of the 65th codon is high. Neither semi-synthetic suppressor tRNA_{CUA} nor any natural tRNA allowed reading of the (*iso*-C)AG codon, as shown by the absence of detectable full-length product in translation mixtures containing the non-standard mRNA in the absence of charged tRNA_{CU(*iso*-dG)}. When the ribosome encountered the (*iso*-C)AG codon in the absence of charged tRNA_{CU(*iso*-dG)}, the primary outcome was continued translation following a frameshift¹⁴ that skipped the *iso*-C base (Table 2). In contrast, attempted translation of a message containing the UAG nonsense codon in the absence of its corresponding charged tRNA_{CUA} resulted in termination at this position, yielding a truncated peptide; there was no detectable frameshifting.

The data in Tables 1 and 2 illustrate the promise of technology that exploits non-standard nucleosides for expanding the genetic lexicon. In addition, they provide a small but important scientific insight suggesting how translation terminates, one of the least

FIG. 1 The mRNA molecules used to compare rearranging (A) and expanding (B) the genetic code with the 65th codon as alternative strategies for incorporating non-standard amino acids into translated peptides. Shown are the translation products arising in the absence (a, octamer) and presence (b, hexadecamer) of suppression, and (c, a dodecamer) following a frame shift and successful termination at the next stop codon (iTy, L-3-iodotyrosine).

A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	AUG	GGU	UUA	UAU	UUG	GGC	CUU	UUU	UAG	GGA	CUC	UAC	CUA	GGG	CUG	UUC	UAA	UGA
	a	Met	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	End								
	b	Met	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	iTy	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	End
B	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	AUG	GGU	UUA	UAU	UUG	GGC	CUU	UUU	iCAG	GGA	CUC	UAC	CUA	GGG	CUG	UUC	UAA	UGA
	a	Met	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	End								
	b	Met	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	iTy	Gly	Leu	Tyr	Leu	Gly	Leu	Phe	End
c	Met	Gly	Leu	Tyr	Leu	Gly	Leu	Phe		Arg	Asp	Cys	Thr	End				

TABLE 1 Label in precipitated products of translation

mRNA	tRNA	Charge	^3H (d.p.m.)	^{35}S (d.p.m.)	Leu:Met	Suppression (%)
1 UAG	none	—	112,979 134,557 106,506	57,673 67,132 50,561	3.0	—
2 UAG	CUA	none	130,382 128,061 120,183	63,814 64,307 59,128	3.0	—
3 UAG	CU(<i>iso</i> -dG)	none	115,937 126,651 106,834	56,437 61,518 53,890	3.0	—
4 UAG	CUA	iodo-Tyr	171,434 219,950 158,723	52,732 65,045 47,313	4.9	63
5 UAG	CU(<i>iso</i> -dG)	iodo-Tyr	141,639 123,592 123,500	63,821 54,144 57,346	3.3	10
6 (<i>iso</i> -C)AG	none	—	30,340 31,871 30,670	14,994 15,405 15,323	3.0	—
7 (<i>iso</i> -C)AG	CUA	none	32,008 32,465 34,310	16,066 15,574 16,607	3.0	—
8 (<i>iso</i> -C)AG	CU(<i>iso</i> -dG)	none	30,986 30,159 33,115	14,723 14,849 16,152	3.0	—
9 (<i>iso</i> -C)AG	CUA	iodo-Tyr	35,644 29,756 31,143	17,363 14,181 15,705	3.0	—
10 (<i>iso</i> -C)AG	CU(<i>iso</i> -dG)	iodo-Tyr	235,233 203,033 202,255	60,179 52,841 53,087	5.7	90

A tRNA^{Gly}_{CUN} (ref. 23) altered in sequence in the acceptor stem and anticodon loop regions and charged with iodotyrosine was prepared by ligating a truncated tRNA lacking the 3'-terminal dinucleotide with a dCA dinucleotide 2'(3') acylated with L-iodotyrosine; both of these were prepared by chemical synthesis^{24,25}. This particular tRNA was chosen to make unlikely any proof reading by the 'double sieve' mechanism when charged with any non-standard amino acid^{25,26}. Messages were prepared by ligation of RNA fragments in the presence of DNA splints. Details will be reported elsewhere (in preparation). Reaction mixtures (10 μl) contained rabbit reticulocyte lysate (9 μl), L-[^{35}S]methionine (15 μCi , 29.7 Ci mmol⁻¹), L-[3,4,5- ^3H]leucine (5 μCi , 19.8 Ci mmol⁻¹), mRNA (2.0 μM), L-3-iodotyrosyl-tRNA (20 μM), and MgCl₂ (1.0 mM). The reaction was quenched with water (1 ml) in the presence of carrier peptides (10 μl of a 0.5 mM solution of each peptide in 77% formic acid), and the precipitated hydrophobic peptide products recovered by centrifugation. Peptides arising from termination have a Leu:Met ratio of 3:1; those derived from complete translation to yield a 16-mer have a Leu:Met ratio of 6:1. The relative amount of terminated and full-length peptides can therefore be calculated from the Leu:Met ratio^{8,13}. The hydrophilic peptides resulting from frame shifting are not recovered by this protocol; see Table 2 for their analysis. Codon and anticodon combinations are indicated. Values are from three trials. Counts per minute (c.p.m.) have been converted into disintegrations per minute (d.p.m.) and corrected for background and quenching. Leu:Met ratios were obtained by correcting for specific activities. Suppression efficiencies are obtained from the average suppression of separate runs derived from the ratio of L-[^{35}S]methionine to L-[^3H]leucine divided by the [^{35}S]:[^3H] ratio obtained from the appropriate control. Data from control experiments with radiolabelled iodotyrosine are not shown. No evidence was found to suggest that the uncharged tRNA (2, 3, 7 and 8) was charged in the lysate during the course of the translation experiment. This is probably because (1) the designed tRNA is structurally altered in the acceptor stem region¹³, (2) it contains no post-transcriptional modifications; and (3) it is derived from an *Escherichia coli* sequence. The possibility that significant amounts of translation products arise from initiation at positions other than the AUG start codon is ruled out by: (1) independent experiments measuring $^{35}\text{S}/^{125}\text{I}$ ratios in control reactions (not shown); (2) the absence of additional products containing tritium in the HPLC analysis; (3) the Leu:Met ratios of 3.0 seen in 1, 2, 3, 6, 7, 8 and 9; were translation being initiated in frame at codons following the AUG, these ratios would require that a corresponding amount of the translation initiating at the AUG be terminated before the stop codon; and (4) consistent ratios of ^{35}S in data obtained by precipitation (Table 1), where randomly initiated peptides translated in one of the other two reading frames would not precipitate, to those obtained by HPLC, where all peptides are, in principle, observable (Table 2).

understood parts of the translation process. Our results can be explained most simply by the hypothesis that release factors¹⁵ present in the translation mixture bind to the nonsense codon UAG but not to the 65th codon (*iso*-C)AG (refs 16, 17). Further, the absence of a complementary tRNA is evidently not sufficient for proper termination of translation; apparently, release factors must bind to prevent frameshifting. The relative rates of termination mediated by release factors and frameshifting control the expression of many viral proteins¹⁸, the expression of subunits of DNA polymerase¹⁹, and the natural reorganization of the genetic code away from the 'universal' code (as for example in mitochondria)²⁰. The 65th codon offers a new experimental strategy to analyse these systems.

From a technological standpoint, the binding of release factors is also critical. Both the genetic code rearrangement and expansion strategies suffer from an intrinsic disadvantage: the low yields of cell-free translation. The presence of release factors in translation mixtures and their requirement for correct termination create a further limitation. To obtain high yields of

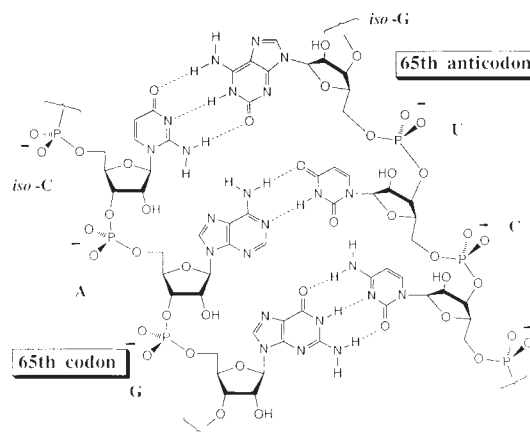


FIG. 2 The 65th codon (incorporating the non-standard nucleoside *iso*-C) and its complementary anticodon (incorporating *iso*-dG) allows the incorporation of non-standard amino acids into proteins synthesized by translation.

TABLE 2 HPLC analysis of the oligopeptides produced by translation

		Full-length products (16-mer)		Termination products (8-mer)		Frame shift products (polar)		
tRNA	Charge	³⁵ S (d.p.m.)	%	³⁵ S (d.p.m.)	%	³⁵ S (d.p.m.)	%	
mRNA containing the UAG codon								
1	none	—	799	4	20,548	96	76	—
2	CUA	none	821	4	19,291	96	29	—
3	CU(<i>iso</i> -dG)	none	524	3	17,844	97	73	—
4	CUA	iodo-Tyr	15,747	67	7,617	33	90	—
5	CU(<i>iso</i> -dG)	iodo-Tyr	1,725	9	18,153	91	47	—
mRNA containing the (<i>iso</i> -C)AG codon								
6	none	—	1,145	3	9,195	25	25,832	71
7	CUA	none	967	4	4,597	17	21,749	80
8	CU(<i>iso</i> -dG)	none	1,023	4	3,400	14	19,078	81
9	CUA	iodo-Tyr	853	3	6,722	24	20,984	73
10	CU(<i>iso</i> -dG)	iodo-Tyr	17,754	91	1,464	8	246	1

Aliquots obtained directly from incubation mixtures were diluted with carrier peptides and injected onto a Vydac C4 column equilibrated in 0.1% trifluoroacetic acid in H₂O/CH₃CN (3:1). Both hydrophobic and hydrophilic peptide, the latter arising from frameshifting, are quantified in this way. Gradient elution with 0.1% trifluoroacetic acid in CH₃CN (25–55% CH₃CN in 60 min) resolved the 8- and 16-mer peaks, eluting at 30–45% CH₃CN. The low read-through of the UAG stop codon seen with non-standard tRNA is expected from the minor tautomeric form of *iso*-G (ref. 27).

translation products, release factors that bind UAG in competition with the semi-synthetic tRNA must be removed or inactivated. However, this would require that UAG triplet stop signals be removed to avoid continued translation after frameshifting. Although it may be possible to fulfil both of these requirements *in vitro*, they are essentially unattainable *in vivo*. Thus for the rearrangement strategy, the low yields of protein obtained from cell-free translation systems cannot be overcome simply by moving the system into a living cell.

Our demonstration that a ribosome can efficiently translate a 65th codon is the first of three breakthroughs required for incorporation of non-standard amino acids into proteins biosynthesized *in vivo*. The second is that plasmids containing a third base pair must be copied and transcribed *in vivo* with reasonable fidelity; the base pair between diaminopyrimidine and xanthosine² may be best suited in this regard. Given faithful replication and transcription, a total of 216 codons will be accessible as a result of the introduction of a third base pair. The third and more demanding requirement is for non-standard aminoacyl tRNA synthetases to be engineered which specifically couple a non-standard amino acid to a non-standard tRNA^{21,22} □

Received 7 November 1991; accepted 13 February 1992.

- Switzer, C. Y., Moroney, S. E. & Benner, S. A. *J. Am. chem. Soc.* **111**, 8322–8323 (1989).
- Picirilli, J. A., Krauch, T., Moroney, S. E. & Benner, S. A. *Nature* **343**, 33–37 (1990).
- Bain, J. D., Glabe, C. G., Dix, T. A., Chamberlin, A. R. & D'Ala, E. S. *J. Am. chem. Soc.* **111**, 8013–8014 (1989).
- Bain, J. D., Wacker, D. A., Kuo, E. E. & Chamberlin, A. R. *Tetrahedron* **47**, 2389–2400 (1991).
- Hecht, S. M., Alford, B. L., Kuroda, Y. & Kitano, S. *J. Biol. Chem.* **253**, 4517–4520 (1978).
- Roesler, J. R., Chorghade, M. S. & Hecht, S. M. *Biochemistry* **25**, 6361–6365 (1986).
- Baldini, G., Martoglio, B., Schachenmann, A., Zugliani, C. & Brunner, J. *Biochemistry* **27**, 7951–7959 (1988).
- Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. *Science* **244**, 182–188 (1989).
- Robertson, S. A., Ellman, J. A. & Schultz, P. G. *J. Am. chem. Soc.* **113**, 2722–2729 (1991).
- Kozak, M. *Nature* **308**, 241–246 (1984).
- Mantsch, H. H. *et al. Biochemistry* **14**, 5593–5601 (1975).
- McBride, L. J., Kierzek, R., Beaucage, S. L. & Caruthers, M. H. *J. Am. chem. Soc.* **108**, 2040–2048 (1986).
- Bain, J. D. *et al. Biochemistry* **30**, 5411–5421 (1991).
- Fluck, M. M., Salser, W. & Epstein, R. H. *Molec. gen. Genet.* **151**, 137–149 (1977).
- Konecki, D. S., Aune, K. C., Tate, W. & Caskey, C. T. *J. Biol. Chem.* **252**, 4514–4520 (1977).
- Smrt, J., Kemper, W., Caskey, T. & Nirenberg, M. *J. Biol. Chem.* **245**, 2753–2757 (1970).
- Valle, R. P. C. & Morch, M.-D. *FEBS Lett.* **235**, 1–15 (1988).
- Craig, W. J. & Caskey, C. T. *Cell* **50**, 1–2 (1987).
- Tsuchihashi, Z. & Kornberg, A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2516–2520 (1990).
- Jukes, T. J., Bessho, Y., Ohama, T. & Osawa, S. *Nature* **352**, 575 (1991).
- Fersht, A. R. *Biochemistry* **26**, 8031–8037 (1987).
- Kast, P. thesis ETH, Zurich (1991).
- Roberts, J. W. & Carbone, J. *J. Biol. Chem.* **250**, 5530–5541 (1975).
- Lyttle, M. H., Wright, P. B., Sinha, N. D., Bain, J. D. & Chamberlin, A. R. *J. Org. Chem.* **56**, 4608–4615 (1991).
- Bain, J. D., Wacker, D. A., Kuo, E. E., Lyttle, M. H. & Chamberlin, A. R. *J. Org. Chem.* **56**, 4615–4625 (1991).
- Fersht, A. R. & Dingwall, C. *Biochemistry* **18**, 2627–2631 (1979).
- Seipol, J., Kazimierzczuk, Z. & Shugar, D. *Z. Naturf.* **31c**, 361–370 (1976).

ACKNOWLEDGEMENTS. We thank S. Hecht and J. Knowles for comments on the manuscript. This work was supported by the Swiss Nationalfond, Sandoz, the ETH and the NIH. C.S. was supported by a fellowship from the NSF.

The dead-end elimination theorem and its use in protein side-chain positioning

Johan Desmet*, Marc De Maeyer†, Bart Hazes‡ & Ignace Lasters†

* Interdisciplinary Research Center, KU Leuven Campus Kortrijk, B8500 Kortrijk, Belgium

† Corvas International NV, Jozef Plateastraat 22, B9000 Gent, Belgium

‡ BIOSON Research Institute, Department of Chemistry, Chemical Physics, RU Groningen, Nijenborgh 16, 9747 AG Groningen, The Netherlands

THE prediction of a protein's tertiary structure is still a considerable problem because the huge amount of possible conformational space¹ makes it computationally difficult. With regard to side-chain modelling, a solution has been attempted by the grouping of side-chain conformations into representative sets of rotamers^{2–5}. Nonetheless, an exhaustive combinatorial search is still limited to carefully identified packing units^{5,6} containing a limited number of residues. For larger systems other strategies had to be developed, such as the Monte Carlo Procedure^{6,7} and the genetic algorithm and clustering approach⁸. Here we present a theorem, referred to as the 'dead-end elimination' theorem, which imposes a suitable condition to identify rotamers that cannot be members of the global minimum energy conformation. Application of this theorem effectively controls the computational explosion of the rotamer combinatorial problem, thereby allowing the determination of the global minimum energy conformation of a large collection of side chains.

The potential energy of a protein system consisting of a set of residue side chains *i* each in a particular rotameric state *r* and embedded in a template of fixed atoms (for example, main-chain atoms) can be written as

$$E_{\text{global}} = E_{\text{template}} + \sum_i E(i_r) + \sum_i \sum_j E(i_r, j_s); \quad i < j \quad (1)$$

where E_{template} is the template self energy, $E(i_r)$ the potential energy of the side chain atoms of the rotamer *i_r* in the force field of the template, and $E(i_r, j_s)$ the nonbonded pairwise interaction energy between rotamers *i_r* and *j_s*. The calculation of the terms $E(i_r)$ and $E(i_r, j_s)$ is computationally quite feasible as their numbers grow linearly and quadratically, respectively, as a function of the number of residues. On the basis of these energies, a condition can be formalized, allowing identification

of rotamers i , that are absolutely incompatible with the global minimum energy conformation (GMEC). In a combinatorial search to identify the GMEC, conformations comprising such rotamers can be qualified as dead ending. Therefore, this condition will be referred to as the dead-end elimination (DEE) theorem: if for a couple of rotamers (i_r, i_i) the following inequality holds true

$$E(i_r) + \sum_j \min_s E(i_r j_s) > E(i_i) + \sum_j \max_s E(i_i j_s); \quad i \neq j \quad (2)$$

then i_r is incompatible with the GMEC. In equation (2) the

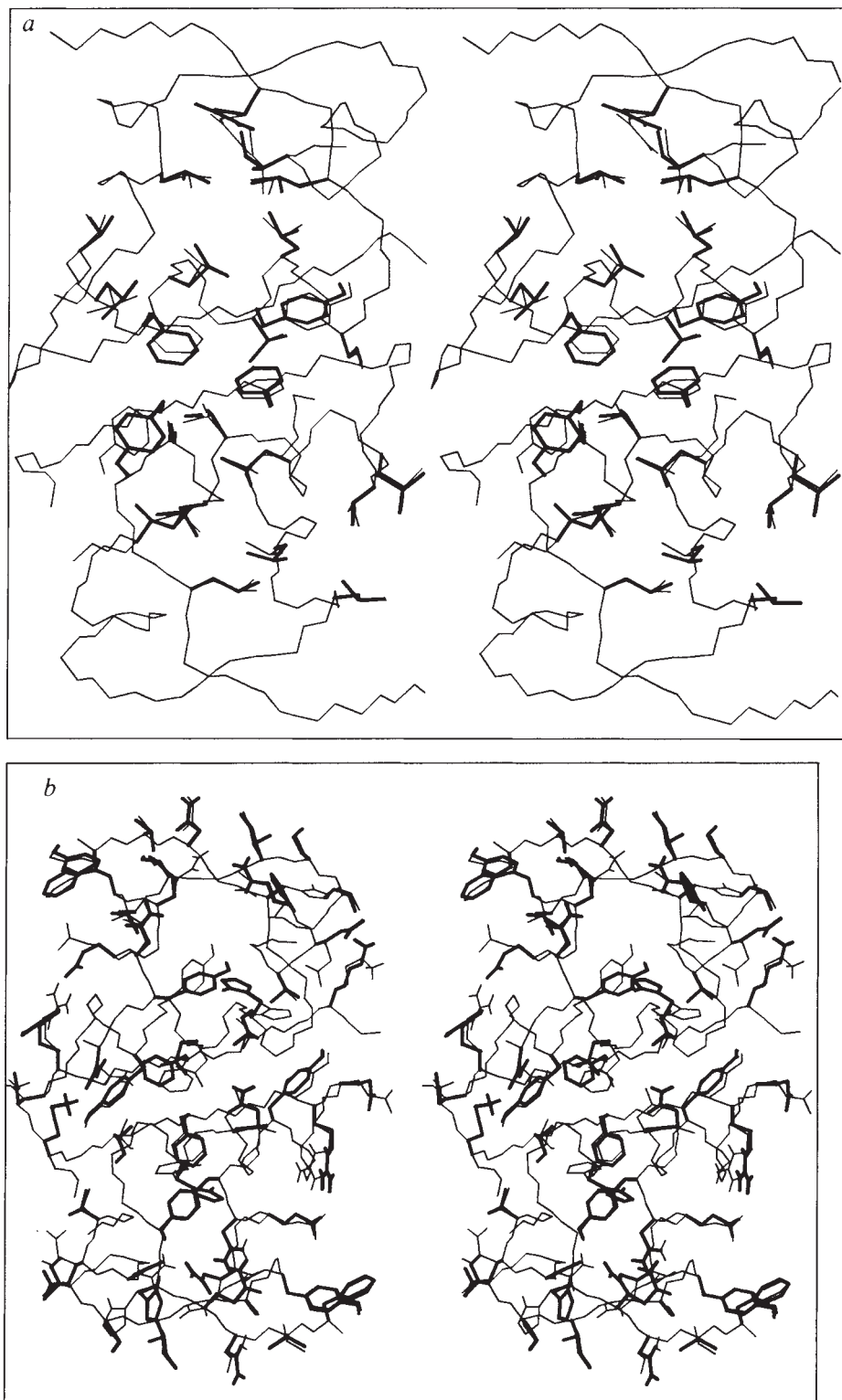
functions $\min_s$ and $\max_s$ are evaluated by searching the lowest and the highest value, respectively, for their argument by cycling over s . This search procedure grows quadratically with the total number of rotamers considered.

The mathematical proof of this theorem readily follows from the definition of the GMEC, which implies that any conformation must necessarily have a higher or equal (in case of degeneracy) energy than the GMEC itself

$$E_{\text{global}} \geq E_{\text{GMEC}} \quad (3)$$

Expressing the right-hand side of this equation making use of

FIG. 1 Stereoscopic view of the insulin dimer (coordinates are obtained from the Brookhaven protein databank¹¹, code 3ins). The other two dimers in the insulin hexamer were included in the prediction procedure to allow a correct modelling of subunit-subunit interfaces. One dimer contains 102 amino acids, 26 of which are Gly, Ala, Pro or Cys involved in a disulphide bridge. These residues do not contain freely rotatable heavy atoms, and were therefore considered as being part of the template. The other 76 residues were simultaneously repositioned. An extended version of the rotamer library⁵ was used. For all aromatic residues, rotamers created by taking steps of size equal to the standard deviation⁵ above and below dihedral angles χ_1 and χ_2 were added to the library. In this way the library, which is considerably larger than those used in previous studies^{5,7,8}, contained 363 rotamers to model 16 residue types. This library is available on request. The number of rotamers for each of the rotatable side chains are Val and Asp, 3; Leu, 4; Ile, 5; Asn, 6; Glu, 7; Ser and Thr, 9; Gln, 10; Met, 21; Phe and Tyr, 36; Lys, 51; His and Trp, 54; Arg, 55. Side-chain placement was done in the presence of all H atoms using a molecular mechanics model¹². The backbone of the complete dimer as well as the residues exposing less (a) or more (b) than 5% of their maximum possible accessible surface area are shown. The predicted conformations are represented in bold lines. To prevent any bias masking the outcome of our method, the modelled structure was not energy-minimized. The r.m.s. deviation between the observed and the predicted heavy side-chain atoms (all residues excluding Gly, Ala, Pro and Cys residues) is 1.69 Å (76 residues). The r.m.s. value calculated with the same criteria for the less than 5% and more than 5% exposed residues are 0.61 Å (24 residues) and 1.96 Å (52 residues), respectively.



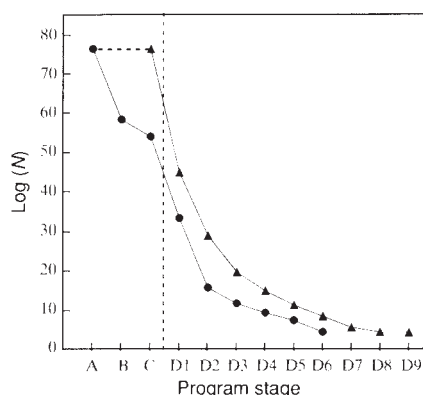


FIG. 2 Evolution of the logarithm of total number of possible rotamer combinations (N) as a function of the program stage in positioning the side chains of the insulin dimer. Program states are as follows: A, start of the program; B, after elimination of rotamers that are clashing with the surrounding template using a threshold energy value of 30 kcal mol^{-1} ; C, after elimination of rotamers that are clashing with all rotamers of some residue using the same threshold; D1–D9, iterations of the DEE procedure. $\blacktriangle$, Evolution of N as a function of program stage, excluding stages B and C. The total required computing time was 5.8 h on an IRIS 4D/31.0 computer of which 5.3 h was spent in the computation of the pairwise rotamer interaction energies (780,000 pairs). The DEE calculations took only 14 min. $\bullet$, Evolution of N as a function of program stage including stages A and B. The total computing time required was 1.8 h of which 1.2 h was consumed in rotamer interaction computation and only about 3 min in DEE calculations.

equation (1), isolating the terms for a particular residue of interest i and denoting GMEC-compatible rotamers with the subscript g , we obtain

$$E_{\text{GMEC}} = E_{\text{template}} + E(i_g) + \sum_j E(i_g j_g) + \sum_j E(j_g) + \sum_{j,k} E(j_g k_g); \quad j, k \neq i; \quad j < k \quad (4)$$

If the residue of interest, i , adopts an undefined rotamer conformation i_r , eventually different from i_g , we can write an analogous expression for the energy of the resulting conformation

$$E_{\text{global}} = E_{\text{template}} + E(i_r) + \sum_j E(i_r j_g) + \sum_j E(j_g) + \sum_{j,k} E(j_g k_g); \quad j, k \neq i; \quad j < k \quad (5)$$

Substitution of equations (4) and (5) into equation (3) gives

$$E(i_r) + \sum_j E(i_r j_g) \geq E(i_g) + \sum_j E(i_g j_g); \quad i \neq j \quad (6)$$

Since for all i, j ($i \neq j$),

$$\max_s E(i_r j_s) \geq E(i_r j_g) \quad (7)$$

and

$$\min_s E(i_g j_s) \leq E(i_g j_g) \quad (8)$$

it follows that

$$E(i_r) + \sum_j \max_s E(i_r j_s) \geq E(i_g) + \sum_j \min_s E(i_g j_s); \quad i \neq j \quad (9)$$

Now, each rotamer i_r that satisfies the dead-end criterion of equation (2) must, in view of equation (9), necessarily obey the inequality

$$E(i_r) + \sum_j \min_s E(i_r j_s) > E(i_g) + \sum_j \min_s E(i_g j_s); \quad i \neq j \quad (10)$$

This implies that $i_r \neq i_g$ thereby proving the DEE theorem.

By the same reasoning, the DEE theorem can be extended to rotamer pairs or, in general, to multi-residue rotamer combinations. Hereafter, only the extension towards pairs is given. The latter is analogous to the single rotamer criterion, and so is the extension towards multiple rotamer entities. Defining a specific energy term for the rotamer pair $[i_r j_s]$, $\varepsilon([i_r j_s])$ as

$$\varepsilon([i_r j_s]) = E(i_r) + E(j_s) + E(i_r j_s); \quad i \neq j \quad (11)$$

and the interaction energy between this pair and another rotamer k_t , $\varepsilon([i_r j_s] k_t)$ as

$$\varepsilon([i_r j_s] k_t) = E(i_r k_t) + E(j_s k_t); \quad i, j \neq k \quad (12)$$

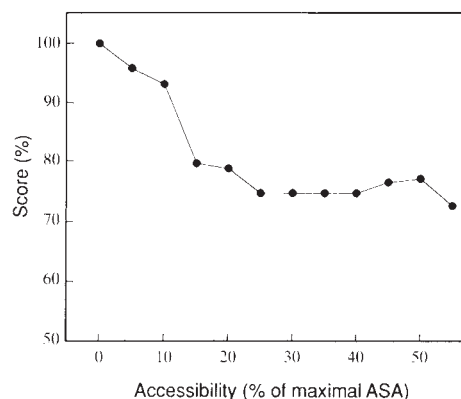
then, a rotamer pair $[i_r j_s]$ is dead ending if there exists a $[i_u j_v]$ pair for which

$$\varepsilon([i_r j_s]) + \sum_k \min_t \varepsilon([i_r j_s] k_t) > \varepsilon([i_u j_v]) + \sum_k \max_t \varepsilon([i_u j_v] k_t); \quad i, j \neq k \quad (13)$$

The proof for this extended DEE theorem can be derived by analogy to the proof of the DEE theorem for single rotamers. Application of the DEE theorem to rotamer pairs grows cubically with the number of rotamers and is therefore the slowest step of the dead-end elimination procedure, once the terms $E(i_r)$ and $E(i_r j_s)$ have been calculated.

The DEE theorem was implemented in a combinatorial substitution algorithm developed in the Brugel⁹ package. We report here an in depth study for the insulin dimer using an extended version of the Ponder and Richards rotamer library⁵ (Fig. 1). The initial number of rotamer combinations (N) for the 76 residues to be positioned in the insulin dimer equals 2.7×10^{76} (Fig. 2, stage A). The power of the DEE theorem is exemplified in the major reduction of this huge rotamer conformational space, especially when applied iteratively (Fig. 2, stages D1 to

FIG. 3 Score for correct prediction of residue side-chain orientation for residues having an accessibility below the indicated value (expressed as per cent). Residues are considered to be correctly predicted if both their χ_1 and χ_2 angles fall in the same dihedral class as observed for the X-ray structure. The accessibility of a residue side chain is defined as the ratio of the ASA in the protein versus the maximal ASA in an extended conformation in a tripeptide and expressed in per cent.



D9). In each iteration stage, two operations are carried out: (1) the DEE theorem is applied to single rotamers until no more dead ending rotamer can be identified; (2) the generalized DEE theorem (equation (13)) is then applied to all possible rotamer pairs and dead ending pairs are flagged. In the next iterative stage, these flagged dead ending pairs can lead to the further elimination of rotamers because: (1) pairs that have been flagged are discarded from the $\sum_j \min_i E(i, j_s)$ computation (equation (2)), which possibly augments the left-hand side of the DEE criterion; and (2) if for a given rotamer all its rotamer pairs with some other residue have been flagged, this rotamer evidently will be dead ending too. The iteration ends when no dead ending rotamers or pairs can be identified. For the insulin dimer this resulted in nine stages (Fig. 2, D1 to D9, upper curve) whereby N was reduced from 2.7×10^{76} to only 7,200, a reduction of N by 72 orders of magnitude.

The number of computations can be considerably reduced by first eliminating rotamers that are clearly incompatible with the given template or with all rotamers of some other residue. Using a 30 kcal mol⁻¹ threshold value, 44% and 6% of the rotamers were eliminated because of incompatibility with template and surrounding side chains, respectively. This elimination reduced the initial N from 2.7×10^{76} to 6.2×10^{53} (Fig. 2, stages B and C). Applying the DEE theorem further reduced N to 10,800 in six iteration cycles (Fig. 2).

After applying the DEE theorem in both experiments, only seven residues had more than one rotamer left. Their side chains were modelled by the following 'add on' mechanism. In each addition step, the rotamers of the added residue were combined with the group of allowed rotameric combinations from the previous step, eventually leading to the rejection of combinations by applying the generalized DEE theorem for multi-residue systems. At each step no more than 23 rotameric combinations had to be stored in memory. The lowest energy conformation obtained in the final addition step was retained as the ultimate structure and compared with the X-ray-determined structure as is illustrated in Fig. 1. Both experiments yielded the same final structure. This must be the absolute GMCC for the given rotamer library and energy function, because in the first experiment no user-defined threshold value was used and also the DEE theorem can only identify GMCC-incompatible rotamers.

Residues are considered to be correctly predicted if both their χ_1 and χ_2 angles fall in the same dihedral class as observed in the X-ray structure. According to this criterion, 55 out of the 76 residues studied were correctly predicted, with average deviations to the X-ray model of $\Delta\chi_1 = 8.9^\circ \pm 7.0^\circ$ and $\Delta\chi_2 = 12.8^\circ \pm 9.2^\circ$. A clear correlation was observed between the correctness of the prediction and the extent of burying (Figs 1 and 3). The residues with a solvent accessible surface area (ASA) up to 10% of their maximum ASA (29 cases) are predicted with an accuracy of 93%. This sharply drops (Fig. 3) to a rather constant level of about 72% when including residues with higher accessibilities. The occurrence of this drop at the level of 10% nicely corresponds to the 5–8% value proposed by Miller *et al.*¹⁰ in their definition of buried residues.

Finally, to assess the prediction power of the DEE theorem for larger systems, a blind test was done on the third domain of Limulus polyphemus subunit type II haemocyanin (resolution = 2.2 Å, R factor = 17.4%). Only backbone coordinates of this structure were provided by one of the authors (B.H., manuscript in preparation). We predicted 194 amino-acid side chains and then compared them with the X-ray model. An overall prediction score of 71% was obtained, with average deviations on χ_1 and χ_2 angles for the correctly predicted side chains of $\Delta\chi_1 = 7.6^\circ \pm 6.5^\circ$ and $\Delta\chi_2 = 15.0^\circ \pm 14.1^\circ$, respectively. Visual inspection suggested that many of the errors were caused by incorrectly placed salt bridges and hydrogen-bonded interactions involving

exposed polar groups. The lower score for these residues can be reasonably attributed to reduced packing constraints and to the absence of water molecules. Modification of the force field is therefore expected to improve the predictions.

We have proposed a theorem that allows, for a given library of rotamers and energy function, an effective search for the GMCC of a large collection of side chains embedded in a given fixed template. In addition to its theoretical interest, we believe that the DEE theorem can be valuable in homology modelling and in predicting side-chain rearrangements, provided a reasonably accurate template is available. Moreover, the high performance of the algorithm now brings into reach intriguing applications, such as flexible substrate-protein docking (whereby protein side chains are allowed to adapt to the substrate molecule being docked) and protein modelling involving main-chain adjustments. □

Received 25 November 1991; accepted 3 February 1992.

- Levinthal, C. *J. chim. Phys.* **65**, 44–45 (1968).
- McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. *J. molec. Biol.* **198**, 295–310 (1987).
- Janin, J., Wodak, S., Levitt, M. & Maigret, B. *J. molec. Biol.* **125**, 357–386 (1978).
- James, M. N. G. & Sielecki, A. R. *J. molec. Biol.* **163**, 299–361 (1983).
- Ponder, J. W. & Richards, F. M. *J. molec. Biol.* **193**, 775–791 (1987).
- Lee, C. & Levitt, M. *Nature* **352**, 448–451 (1991).
- Holm, L. & Sander, C. *J. molec. Biol.* **218**, 183–194 (1991).
- Tuffery, P., Etchebest, C., Hazout, S. & Lavery, R. *J. biomolec. struct. Dyn.* **8**, 1267–1289 (1991).
- Delhaise, P., Bardiaux, M. & Wodak, S. *J. molec. Graph.* **2**, 103–106 (1984).
- Miller, S., Janin, J., Lesk, A. M. & Chothia, C. *J. molec. Biol.* **196**, 641–656 (1987).
- Berenstein, F. C. *et al. J. molec. Biol.* **112**, 535–542 (1977).
- Brooks, B. R. & Karplus, M. *J. comput. Chem.* **4**, 187–217 (1983).

ACKNOWLEDGEMENTS. We thank W. Hol for discussion. J.D. thanks F. Van Cauwelaert for scientific advice, S. Wodak for introduction to the modelling field and the Research Foundation of Leuven and the FGWO for financial support.

GUIDE TO AUTHORS

Please follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributions should therefore be written in English. Manuscripts in other languages and those for which English is a second language, and written clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruption of the text, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest referees; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, who are chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive galley proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check their proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned, but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items (figures and tables) and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarizes its results and implications. Articles have fewer than 50 references and may contain a few subheadings of two or three words.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the commentary editor in the form of a one-page synopsis. Commentaries are normally between one and four pages of *Nature*.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the

News and Views editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words and 5 references.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and two copies are required, each accompanied by artwork. If photographs are included, four sets of originals are required: for line drawings, one set of originals and three good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Revised and resubmitted manuscripts should also be clearly marked as such and labelled with their manuscript numbers.

Titles say what the paper is about with the minimum of technical terminology and in fewer than 80 characters in the case of Articles and Letters. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Abstracts should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided, and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs. Numbering of sequences should be in the left-hand margin only, with a single space between rows.

Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent the publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. References to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and entered in the reference list. All other forms of reference, including unlettered abstracts, should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used except for changed addresses.

Acknowledgements are brief; grant and contribution numbers are not allowed.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1137 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air to the Editor to London should be declared as manuscript and value \$5 to prevent the imposition of import duty and value-added tax (VAT). □

Protein model building using structural homology

Richard H. Lee

Homology modelling software can predict the structure of a newly sequenced protein by incorporating conformational similarity information from related proteins for which atomic coordinates are known. Models built using structural data are more reliable than those based on sequence alignments alone.

PROTEIN and nucleic acid sequences are being published at an ever-increasing rate. There are, for example, over 20,000 protein sequences listed in the database distributed by the National Biomedical Research Foundation¹. At the same time, structure determination by X-ray crystallography and NMR methods is done at a much slower rate. The result is that there is a whole range of proteins for which the sequences are known, but the structures are not.

Homology modelling is a technique that can be used to exploit the structural and sequence similarities among members of a protein family to predict the conformation of a newly characterized protein. It is more computationally efficient than such conformational searching methods as molecular dynamics or Monte Carlo sampling. The information required to build a model protein consists of its sequence and the atomic coordinates of at least one other protein. Early studies^{2,3} used only a single reference structure, predicting the coordinates of the model based on the relative similarities of the two sequences alone. As there are instances where two sequences are similar, but the conformations are not, and vice versa, it is much more desirable to utilize more structural information. Thus, in the methods of Greer⁴⁻⁶ and Blundell^{7,8}, the structural similarities between the reference proteins are found, and it is extrapolated that the model protein will have similar conformations in corresponding regions. More confidence can be placed in a model built this way because information from the entire set of related proteins is incorporated.

Seven steps to model building

The process of building a model protein using homology techniques can be divided into seven steps. (1) Search one or more sequence databases^{1,9-12} to find proteins related to the model to be built. (2) If more than one protein structure is available, find those areas that are similar among the reference structures (the structurally conserved regions or SCRs). (3) Find similarities between the sequence of the model protein and those of the reference set of proteins in the conserved regions (sequence alignment). (4) Set the

conformation of the model protein to be the same as that of one of the reference proteins in each region. Care must be taken to choose appropriate side-chain conformations. (5) Determine the conformations of the variable regions through structural database searching¹³ or loop generation¹⁴. (6) Complete the model by supplying coordinates for the N- and C-terminal residues. (7) Refine the model through energy minimization and molecular dynamics to relieve the strain caused by misplaced side chains, inappropriate peptide chain topologies in the loops, and so on.

Modelling software packages, such as the Homology program offered by Biosym Technologies (San Diego, California), provide facilities to accomplish each of these steps. The remainder of this article will focus on step 2.

As stated, if more than one reference structure is available, it is advisable to find those parts of the proteins that have similar conformations. This can be time-consuming if done manually. Methods to automate the process have been suggested by Taylor and Orengo¹⁵⁻¹⁷ and Rose *et al.*¹⁸. A new algorithm is described here to identify automatically structural conservation in two proteins. A list of matching residues is produced, comprising the SCRs, and the sequences are aligned to reflect the results.

For each of the two proteins being compared, a C_{α} distance matrix is constructed. A distance matrix of this type contains all the conformational information necessary and has the added advantage that it is independent of any coordinate frame. In general, if two distance matrices are found to be the same, then the molecules from which they were derived must also be the same, except for a possible difference in chirality.

A small portion of the matrix of one of the proteins is selected to be compared to that of the other protein. As SCRs are defined as contiguous peptide segments,

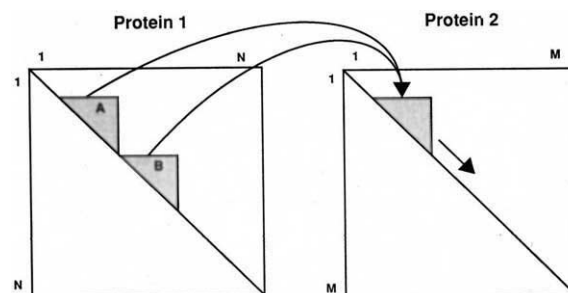


FIG 1. Assessing peptide segment similarities. A probe from the distance matrix of Protein 1 is overlaid onto that of Protein 2, and the RMS difference of the corresponding matrix elements is calculated. The probe is then moved down the diagonal of the second matrix, and the calculation is repeated. A minimum value, which represents the best match of the peptide segment in Protein 1 to a segment in Protein 2, is saved if the RMS difference is below a specified threshold. The entire process is then repeated with a new probe from Protein 1.

matrix elements adjacent to the diagonal are chosen as the probe. The probe, which is chosen to have a fixed length, is overlaid onto the matrix of the other protein, and a root mean squared (RMS) difference of the corresponding matrix elements is calculated. Then the probe is shifted down along the diagonal of the distance matrix of the second protein, and the calculation is repeated (see Fig. 1).

The relative position of the probe along the diagonal when the RMS difference is at a minimum defines two peptide segments that are conformationally similar, at least locally. A threshold is specified to define how small the RMS difference has to be before two segments are considered to be similar. This entire procedure is repeated for a series of probe segments chosen from the first protein.

Two peptide segments cannot be considered to be similar for a homology modelling project unless they both share the same orientation with the rest of their respective proteins. If their orientations are different, then, in general, it is not possible to decide which orientation is correct. Therefore, it is necessary to test this characteristic as well. For a pair of segments (represented by the corresponding matrix elements) found in the preceding step, the off-diagonal matrix elements are compared. In Fig. 2, the peptide segment represented by probe A has been

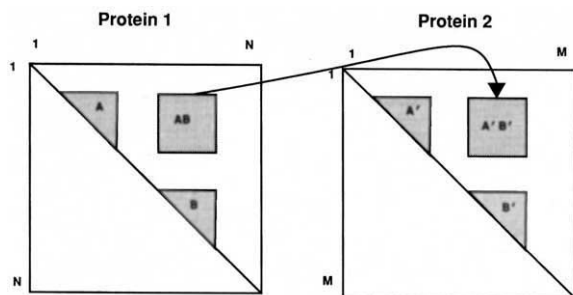


FIG. 2 Checking the relative orientations of pairs of segments. In order to build a reliable model, the relative orientations of all the structurally conserved regions must be similar. Information about the orientation of two segments is contained in the off-diagonal block of matrix elements. If the RMS difference of the matrix elements in the off-diagonal blocks of two proteins is small, then the two segments are oriented similarly in both proteins. Thus, if segment A corresponds to segment A', and segment B corresponds to segment B', then if the RMS difference of AB and A'B' is below a specified threshold, then A and B have the same relative orientation as A' and B'.

found to be locally similar to A', and segment B has been matched to B'. The individual orientations are described in the off-diagonal blocks AB and A'B'. The RMS difference between the two matrix element blocks is calculated; if this value is below a second threshold, then the two peptide segments are said to have similar relative orientations.

The segments found in the first step are then examined for self-consistency. A matrix is constructed containing the orientation scores for all possible pairs. A segment is retained only if the RMS differences of the off-diagonal matrix element blocks between itself and more than half of the other segments are below the threshold. This restriction is necessary to eliminate sets of segments that are consistent with each other, but are not consistent with the major part of the protein.

Many of the segments are overlapping. The final set of SCRs is found by scanning the lengths of the two sequences and combining overlapping segments, stopping when one or more residues do not have a match in the other protein.

When defining SCRs in reference proteins for a model building project, it is best to terminate arbitrarily each region at the end of a secondary structural unit (helix or single strand) and not to include a turn in an SCR. This is done to allow for insertions and deletions at the turns, which generally occur at the surface of the protein. As written, the algorithm cannot distinguish secondary structure, and so many of the SCRs that were found would need to be divided into shorter ones before coordinate assignment is done.

The process of building a model protein through homology techniques is a fairly straightforward one. The accuracy of the final result is dependent entirely on the quality of the information brought to the project in the form of sequence alignments and structural data. Modelling programs such as Biosym's Homology can help by providing convenient tools and automated procedures to enable the researcher to predict several candidate structures and compare their merits, thus making an interactive approach to model building feasible. □

Finally, the sequences are aligned to reflect the matches found between the residues.

Practical applications

The algorithm was used to determine the areas that are structurally conserved between the serine proteases trypsinogen and elastase. Ten SCRs were found, ranging in length from 9 to 34 residue pairs. The overall RMS deviation of the backbone atoms for all SCRs combined was 1.02 Å. All conserved secondary structure elements were included in one or another SCR, and insertions and deletions were clearly identified (see Fig. 3).

Richard H. Lee is manager of the Homology program at Biosym Technologies, Inc., 9685 Scranton Road, San Diego, California 92121, USA. The work was supported in part by Small Business Innovation Research grants from the National Science Foundation. For more information, fill in reader service number 100.

1. NBRF/PIR: National Biomedical Research Foundation, 3900 Reservoir Road, NW Washington, DC 20007, U.S.A.
2. Browne, W. J. *et al. J. molec. Biol.* **42**, 65-86 (1969).
3. Shotton, D. M. *et al. Nature* **225**, 811-816 (1970).
4. Greer, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3393-3397 (1980).
5. Greer, J. *J. molec. Biol.* **153**, 1027-1042 (1981).
6. Greer, J. *Science* **228**, 1055-1060 (1985).
7. Blundell, T. L. *et al. Nature* **326**, 347-352 (1987).
8. Blundell, T. L. *et al. Eur. J. Biochem.* **172**, 513-520 (1988).
9. EMBL database: European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, Germany.
10. GenBank database: IntelliGenetics, Inc., 700 El Camino Real East, Mountain View, CA 94040, U.S.A.
11. Lipman, D. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
12. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444-2448 (1988).
13. Jones, T. A. & Thirup, S. *EMBO J.* **5**, 819-822 (1986).
14. Shenkin, P. S. *et al. Biopolymers* **26**, 2053-2085 (1987).
15. Taylor, W. R. & Orengo, C. A. *J. molec. Biol.* **208**, 1-22 (1989).
16. Taylor, W. R. & Orengo, C. A. *Protein Eng.* **2**, 505-519 (1989).
17. Orengo, C. A. & Taylor, W. R. *J. theoret. Biol.* **147**, 517-551 (1990).
18. Rose, J. & Eisenmenger, F. *J. molec. Evol.* **32**, 340-354 (1991).

ADVERTISEMENTS

**GROW YOUR E. COLI TO
HIGH CELL DENSITIES WITH....**

"TERRIFIC BROTH"

- ☆ STERILE, READY TO USE.
- ☆ MOLECULAR BIOLOGY GRADE.
- ☆ HIGH DENSITY CULTURE OF E. COLI
- ☆ FOR PLASMIDS AND COSMIDS.

Bio Tools™

P.O. BOX 10, PONTYCLUN, MID GLAMORGAN
CF7 9YA. UK FAX/TEL XX-44-(0)443 225574

Reader Service No.11

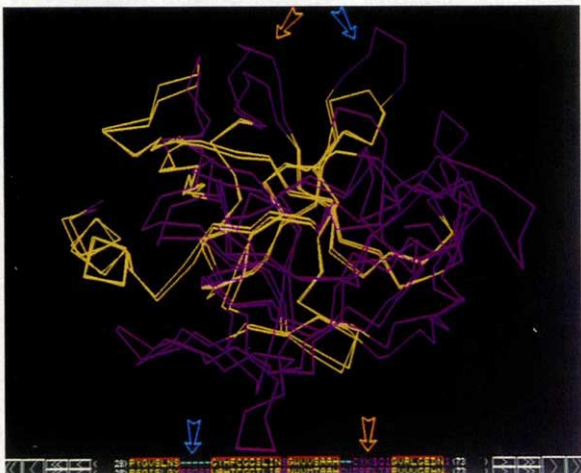


FIG. 3 The two serine proteases trypsinogen and elastase show remarkable structural homology. Ten structurally conserved regions (SCRs) and the corresponding sequence alignment were found automatically using the protein comparison algorithm of the Homology program. The SCRs are shown in yellow, and they were computed to have an overall RMS deviation of the backbone atoms of 1.02 Å. One insertion in each of the two proteins is shown by colour-coded arrows, both in the sequences and in the structures.

Synthetic Oligonucleotides

*Custom made probes and primers,
synthesised, purified and shipped in 3 days*

£3.50 per base
(All inclusive)
0.2 µM scale (approx. 500 µg)

**Modified bases, biotin or digoxigenin labelling,
data base analysis and full design facilities**

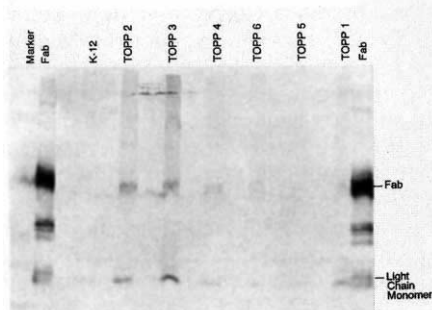
Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX
Tel. XX-44-(0)71 - 326 3126
FAX XX-44-(0)71 - 733 3877

Reader Service No.56

Proteins by design

A competent cell kit for improved protein expression, a new glycan release and preparation system for glycobiology research and a molecular modelling package that turns a desktop PC into a three-dimensional molecular modelling and analysis system — new tools for protein engineering.

STRATAGENE has introduced six strains of TOPP Epicurian Coli **competent cells**, which are designed to overexpress various recombinant proteins known to be difficult or impossible to express in *Escherichia coli* K-12 (*Reader Service No. 101*). Distantly related to *E. coli* K-12,



Expression of Fab antibody in *Escherichia coli* K-12 and TOPP cells — see Stratagene.

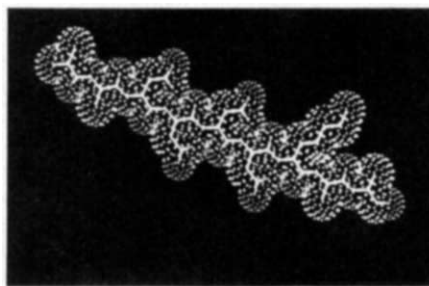
Stratagene says the TOPP *E. coli* cells can provide a rapid, easy and inexpensive way to improve protein expression without sub-cloning, altering growth conditions or changing to a eukaryotic host-vector system. The conventional alternatives can take months to complete, in contrast to the week needed to try TOPP cells. In many cases, Stratagene says that the TOPP strains have dramatically increased the production of recombinant proteins such as human Fab antibody protein and tenascin, a human extracellular-matrix protein. As the TOPP hosts are *E. coli*, commonly used expression vectors like pBluescript II phagemid and pUC can be transformed and propagated in these strains. The TOPP competent cell kit includes all six strains (three 0.2-ml tubes of each). Each strain is also available individually. TOPP cells are the latest addition to the Stratagene line of Epicurian Coli competent cells, which includes XL1-Blue strains, SURE super-competent cells, as well as AG-1, JM101, JM109, JM110, NM522 and SCS-1 strains.

The Nest Group is offering a new chemistry for 'reverse' reverse-phase chromatography of polar peptides (*Reader Service No. 102*). This new chemistry for hydrophilic interaction chromatography or reverse, reverse-phase chromatography, is designed for more polar molecules that exhibit little or no retention under conventional reverse-phase conditions. Polyhydroxyethyl aspartamide uses high organic concentrations

and a sodium perchlorate gradient causing peptides and other polar molecules to exhibit normal phase elution. It is suitable for histones, membrane proteins, phosphorylated amino acids, RGD peptides or peptides that retain poorly on RPC. Nest says that membrane-associated proteins which require high organic concentrations to dissolve may be separable via this method. This column chemistry can also be used for small peptide SEC (mwt 500–10,000) by adjusting the mobile phase with formic acid or phosphate.

Cleavage agents

Ultrapur **enterokinase**, a specific protease available from Biozyme, has been highly purified for the cleavage of fusion protein containing the entero-



Biozyme's ultrapur enterokinase.

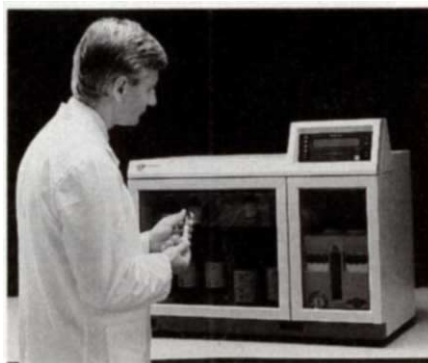
kinase recognition sequence (*Reader Service No. 103*). The enterokinase, which has a high specific activity (>100,000 units per milligram of protein — trypsinogen as the substrate, 25 °C, pH 5.6), can be used in minimal quantities reducing the risk of product degradation. The enzyme is available in both research and bulk production quantities.

With the introduction of chymotrypsin, endoprotease Arg-C and chemically-modified trypsin, Boehringer presents a comprehensive range of **sequencing-grade proteases** (*Reader Service No. 104*). Boehringer says that the preparations are free of contaminating proteolytic activities that might interfere in the separation range of HPLC. And, the proteases are free of proteinous stabilizers, such as bovine serum albumin. According to the company, each lot is tested for its purity and the specificity is verified using defined peptide substrates. The specially-modified trypsin is designed to be more stable with regard to autolysis than native trypsin — even after long incubations of greater than 18 hours, Boehringer says no tryptic peptides are found in the digestion

mix. In addition to the sequencing-grade proteases, Boehringer offers IgA-protease, a restriction protease that cleaves substrate proteins only within or after a typical recognition sequence, which for IgA-protease is as follows: Y-Z-Pro-Pro-/-X-Pro (X = Thr, Ser or Ala; Y = preferentially Pro or Ser; and Z = preferentially Arg or Thr). IgA-protease is provided in a glycerol-containing solution and packaged in aliquots.

System set-ups

The first and historically most difficult step in glycoprotein characterization is to obtain an authentic, pure and intact pool of glycans from the glycoprotein under study. The reactions and chromatographic procedures to achieve this have now been fully automated by Oxford GlycoSystems with the introduction of the GlycoPrep 1000 **glycan release and preparation system** (*Reader Service No. 105*). With the GlycoPrep 1000, new research studies into the biological relevance of glycans are possible. In recombinant glycoprotein development, the GlycoPrep 1000 will make lot-to-lot assessment and cell-line selection economic, easy and reproducible, the company says. The system auto-



Glycobiology research system.

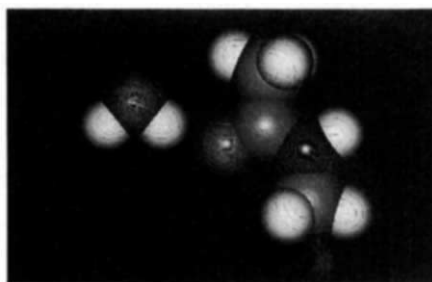
mates hydrazinolysis of up to two sample glycoproteins with conditions optimized for the non-selective release of both N- and O-glycans. A separate "O" mode is available for the selective release and recovery of O-glycans. The pool of intact N- and O-glycans with free reducing termini and in the same molar proportions as in the sample glycoprotein, is presented to the operator after an overnight run.

The Waters PrepLC 2000 **preparative chromatography system** from Millipore is a fully integrated and auto-

mated system for preparative purification and isolation (*Reader Service No. 106*). The system can be used to perform a variety of tasks ranging from the purification of organic synthesis and natural products to the isolation of peptides and proteins. It features gradient solvent delivery, automated sample injection, a 'recycle' feature that allows the user to conserve solvent by recycling difficult-to-separate compounds through the system and to achieve better resolution with each pass, flow rates from 4–300 ml min⁻¹, a 2,000 p.s.i. backpressure capability and manual or automated fraction collection.

Making models

HyperChem from Autodesk is software that integrates powerful molecular modelling and analysis tools into a pack-



HyperChem — Autodesk's 3-D molecular modelling and analysis system. Here, the H₂O and N-methyl-acetamide system was optimized using the AMBER force field, and the total charge density plot of the system was calculated using CNDO.

age that runs on a desktop computer (*Reader Service No. 107*). The software was designed to enable the user to model and analyse three-dimensional (3-D) structures without needing to spend large amounts of time or money. Running under Windows, HyperChem features computational chemistry capabilities previously available only on workstation or mainframe computers, along with a graphical users interface, and a full suite of modelling and visualization tools. HyperChem allows the user to create two-dimensional and 3-D structures easily using a mouse; build protein and nucleic acid structures from amino acids and nucleotide residue libraries; display models as sticks, filled spheres (CPK), dot surfaces or overlapping disks; rotate, translate and align structures along x, y, z axes; perform real-time 3-D solids rotation with advanced display systems such as those from Silicon Graphics; label residues or highlight the backbone of proteins and nucleic acids; obtain or modify structural parameters such as bond lengths, bond angles and torsion angles; and read/write Brookhaven Protein Database files. In addition, users can perform classical mechanics calculations using MM2, AMBER, CHARMM and OPLS force fields and semi-empirical quantum mechanical calculations using Extended Hückel, CNDO, INDO, MINDO/3, MNDO or AM1 methods. Users can display semi-empirical calculation results graphically, superimposed on the molecular structure, use molecular dynamics calculations to simulate behaviour of a molecular system under varying temperature conditions and simulate solvation of molecules with periodic boundary conditions. HyperChem runs on a 386/486 PC or UNIX platform.

Commet is a **molecular modelling package** from Phase Separations comprising software and plug-in transputer boards to fit IBM or compatible PCs (*Reader Service No. 108*). Building fragments of proteins from the built-in library of amino acid residues in any secondary structural conformation should present no difficulties, the company says.

The fold predictor module can be used to predict the locations of secondary structure features from a given amino acid sequence, using either the Robson or Chou and Fasman algorithms. Other features include oligopeptide excision, the substitution, insertion and deletion of residues, annealing capabilities, hydropathy plots, database (Brookhaven Protein Data Bank) searching, docking facilities and a host of display options.

Sequence alignment

MultiAlin is a new **multiple sequence alignment program** from Cherwell Scientific designed for the fast sequence analysis of proteins and nucleic acids on a PC (*Reader Service No. 109*). Developed by Florence Corpet of INRA, Toulouse, France, this new version of MultiAlin comes with a new user interface and a sequence alignment screen editor that allows manual refinement of the analysis. With MultiAlin, users can perform automatic multiple alignment with rapid iteration to a consensus sequence; import sequence data in several formats, including GCG, GenBank, EMBL, NBRF and IntelliGenetics; select a comparison table such as DAYHOFF for amino acids and GENETIQ for nucleic acids; and choose penalties for fixed or variable gaps. Other features of the program include up to 16,000 residues in any one sequence, global alignment of 250 sequences at a time, and scores for the multiple alignment quality. MultiAlin will run on any IBM PC or compatible that runs DOS version 2.0 or later.

Version 2.0 of GeneWorks, a **DNA and protein sequence analysis program** for the Apple Macintosh, is now available from IntelliGenetics (*Reader Service No. 110*). The program offers true multiple sequence alignment, calculating functional and evolutionary relationships between sequences and families of sequences. An editor allows the user to make minor alterations to GeneWorks' high-quality alignments, IntelliGenetics says. GeneWorks 2.0 assembles contigs from overlapping sequencing gels, automatically comparing the complements of all sequences. An editor, which simultaneously displays translations and restriction maps, allows the user to locate and correct sequence ambiguities. Dot matrix similarity plots and PROSITE pattern matching completes the line-up of new features. Database capabilities include boolean searches and fast similarity searches of nucleic acid and protein data banks on CD-ROM. The GeneWorks CD-ROM contains the GenBank and EMBL nucleic acid data banks and the SWISS-PROT protein data bank. GeneWorks runs on a minimum configuration of an SE30. It costs \$2,750 (US) for academic users, \$3,300 (US) for commercial users. □

ADVERTISEMENTS

Your Valuable Protein

MICHAEL WEST CONSULTING

Toronto, Ont.

provides top Quality Service:

- PURIFICATION • PROTEOLYSIS
- PEPTIDE SEPARATION
- SEQUENCING

CONSULTATION
AND TRAINING ON SITE!
PHONE FAX

416-255-9288

416-979-5294

NH₂-Ala-Glu-Asn-Ser.....
YOUR SEQUENCE

Reader Service No.13

Professional Custom Services

DNA SYNTHESIS
PROTEIN SEQUENCING
PEPTIDE SYNTHESIS

Rapid Service • Guaranteed Quality • Ample Quantity
Competitive Prices • Free Quotations
World-wide Delivery

V.G. Services

1200 Wonderland Rd. S. Bldg #7, Unit #2
London, Ontario Canada N6L 1A6

CALL/FAX (519) 652-3758 ALT. FAX (519) 652-9622

Reader Service No.19

PEPTIDE
PRODUCTS LTD.
Custom Synthesis & Specialist
Peptide Chemistry

High quality standard services

Custom peptide synthesis
HF cleavage and purification
Peptide conjugation and antibody production
Synthesis of modified amino acids and peptides

Peptide Products Ltd.

David Hart Business Centre,
Unit 2, Salisbury Road,
Downton, SP5 3HT
Wiltshire, England (U.K.)

Tel: 0725 23021
Fax: 0725 23024

Reader Service No.10